

NAMUR
1960

3

Cybernetica

ASSOCIATION INTERNATIONALE DE CYBERNÉTIQUE
INTERNATIONAL ASSOCIATION FOR CYBERNETICS

Sous la Présidence d'honneur de M. le Gouverneur de la Province de Namur

Conseil d'Administration
Board of Administration

PRÉSIDENT :

M. Georges R. BOULANGER (Belgique), Professeur à la Faculté Polytechnique de Mons et à l'Université Libre de Bruxelles.

MEMBRES :

MM. René CLOSE (Belgique), Avocat.

Louis COUFFIGNAL (France), Inspecteur Général de l'Instruction Publique, Directeur du Laboratoire de Calcul Mécanique de l'Institut Blaise Pascal, Paris.

John DIEBOLD (U.S.A.), President of John Diebold and Associates, Inc., New York.

W. Ross ASHBY (United Kingdom), M.D., D.P.M., Director of the Burden Neurological Institute, Bristol.

ADMINISTRATEUR-DÉLÉGUÉ :

M. Josse LEMAIRE (Belgique), Directeur de l'Office Économique, Social et Culturel de la Province de Namur.

CYBERNETICA

est la revue de l'Association Internationale de Cybernétique.

Elle paraît 4 fois par an.

is the review of the International Association for Cybernetics.

It is issued four times a year.

Prix et conditions de vente — Price and conditions of sale.

Abonnement annuel — *Yearly subscription :*

membres de l'Association	150,- F. B.
<i>members of the Association</i>	150,- F. B.
non-membres :	300,- F. B.
<i>non-members :</i>	300,- F. B.

Par numéro — *Each number :*

membres de l'Association	50,- F. B.
<i>members of the Association</i>	50,- F. B.
non-membres :	100,- F. B.
<i>non-members :</i>	100,- F. B.

Toute correspondance concernant la revue est à adresser à l'Association Internationale de Cybernétique, 13, rue Basse Marcelle, Namur (Belgique).

All correspondence concerning the review is to be sent to the International Association for Cybernetics, 13, rue Basse Marcelle, Namur (Belgium).

Secrétaire de Rédaction : M. Roger DETRY

CYBERNETICA

VOLUME III

N° 3 - 1960

Revue de l'Association Internationale de Cybernétique
Review of the International Association for Cybernetics

NAMUR

Les articles sont rédigés en français ou en anglais au choix de leurs auteurs. Ils n'engagent que ces derniers.

La reproduction intégrale ou abrégée des textes parus dans la revue est interdite sans autorisation spéciale de l'Association Internationale de Cybernétique.


The papers are written in English or in French according to the choice of their authors and on their own responsibility.

The complete or the partial reproduction of the papers printed in the review is forbidden without special authorization of the International Association for Cybernetics.

SOMMAIRE

CONTENTS

H. VALLET : <i>Essai d'une théorie générale de l'information</i>	175
H. LABORIT and B. WEBER : <i>The value of the application of cybernetic principles to physiological regulatory mechanisms</i>	216
M. L. HURNI : <i>Observations on the role of business research as an aid to managers</i>	234



Digitized by the Internet Archive
in 2024

Essai d'une théorie générale de l'information

par Henri VALLET,

Ingénieur Radio E. S. E. (Paris)

Il existe, actuellement, une dissymétrie considérable entre la thermodynamique, théorie traitant de l'énergie sous sa forme la plus générale, et les théories traitant de l'information.

La thermodynamique est l'un des chapitres les mieux établis de la physique. Son objet est l'étude de l'énergie, au sens le plus général (c'est-à-dire indépendamment de la forme particulière sous laquelle elle se présente). L'énergie y obéit à des lois précises et se présente comme une grandeur mesurable et bien définie. Il est possible de passer d'une forme d'énergie à une autre avec un rendement maximum calculable et, surtout, il est possible d'établir une théorie de l'énergie indépendamment de sa transmission. Cette théorie, la thermodynamique, est alors complétée par nombre de théories secondaires qui considèrent une forme particulière d'énergie et étudient son comportement, c'est-à-dire — dans le cas de ces théories secondaires — essentiellement sa transmission.

Rien de semblable dans le cas de l'information. La grandeur elle-même est mal définie. Ses règles de combinaison ne sont pas incontestables. Elles ne paraît pas obéir dans tous les cas à la loi de la conservation, ou tout au moins, cette loi apparaît comme beaucoup moins nettement énoncée que dans le cas de l'énergie. Les diverses formes d'information ne sont pas définies et délimitées clairement et il ne paraît pas y avoir pour l'information, de théorème (identique au théorème de Carnot pour l'énergie) qui impose au rendement une limite théorique lors d'un échange d'information.

Cette dissymétrie nous semble en fait plus apparente que réelle. Elle vient à notre avis, du fait que l'on compare ce que l'on appelle actuellement la « théorie de l'information » à la thermodynamique. La comparaison qui s'imposerait, en fait, se situerait plutôt au niveau d'une théorie particulière (celle de la propagation de la

chaleur, par exemple). La théorie de l'information, sous sa forme actuelle, est, en effet, essentiellement une théorie de la transmission de l'information. Les tentatives pour mettre au point une théorie de l'information, indépendamment de sa transmission, ne paraissent pas avoir connu un plein succès. C'est ce fait qu'Aigrain ⁽¹⁾ présente ainsi :

« Remarquons enfin que l'élément essentiel de la théorie de » Shannon est l'idée de transmission, beaucoup plus que celle » d'information. Le fait, en particulier, que la vitesse d'information » d'un message dépende fortement des caractéristiques de l'utili- » sateur — et cela n'est que bien naturel — permet de douter du » bien-fondé des tentatives répétées de créer une théorie de l'in- » formation indépendamment de ce que l'on désire faire de cette » information ».

Notre intention n'est pas de présenter ici une théorie traitant de tous les aspects de l'information, au sens le plus général. Il s'agirait alors d'un travail considérable. Nous voulons, plus simplement, donner quelques indications permettant le développement d'une théorie générale de l'information, symétrique de la thermodynamique et se situant au même niveau de généralité. La voie dans laquelle nous nous engagerons pour cela n'a pas encore été tentée (du moins à notre connaissance).

Une telle théorie présente en effet un réel intérêt et plusieurs raisons militent en sa faveur :

1^o) L'information apparaît comme absolument inséparable de la notion d'énergie : une transmission d'information est inconcevable sans un support énergétique. L'amplitude de ce support peut être aussi faible qu'on le désire ; elle ne peut jamais être nulle.

2^o) Il est d'ailleurs possible d'affirmer, en contrepartie, qu'aucun échange énergétique ne peut s'effectuer sans entraîner un transfert d'information. Cette information peut être utilisée ou non ; sa possibilité n'en existe pas moins dans tous les cas.

3^o) On peut en déduire qu'il existe une liaison indissoluble entre énergie et information. Personnellement, ces deux termes nous apparaissent en fait comme deux aspects d'une même entité fondamentale, et la limite d'aboutissement d'études sur ce sujet devrait être une théorie unitaire de l'énergie et de l'information.

4^o) L'élaboration d'une telle théorie sera grandement facilitée

(1) *La Cybernétique* (Réunion L. DE BROGLIE), Édition de la Revue d'Optique, 1951.

si l'on met tout d'abord sur un pied d'égalité l'énergie et l'information, en élaborant pour cette dernière une théorie se situant au même niveau de généralité que la thermodynamique.

C'est le but du travail actuel pour lequel nous prendrons ce parallélisme — qui nous paraît fondamental — comme postulat de départ. A partir des principaux résultats de la thermodynamique (il est essentiel de remarquer que nous ne considérerons ici que la thermodynamique macroscopique), nous tenterons de définir les notions correspondantes dans une théorie générale d'information. Nous énoncerons les résultats qui représentent pour l'information le parallèle des principaux résultats de la thermodynamique.

RAPPELS DE THERMODYNAMIQUE

Il nous a paru plus pratique, pour la clarté de l'exposé, de rappeler brièvement les principales notions de thermodynamique, qui, pour être parfaitement classiques, ne sont toutefois pas forcément présentes dans toutes les mémoires sous la forme et dans l'ordre que nous adoptons ici. On pourra ainsi mettre en évidence plus aisément le parallélisme des deux théories.

OBJET DE LA THERMODYNAMIQUE

La thermodynamique englobe toutes les modifications des corps matériels et, dans la plupart de ses raisonnements, considère un système qui subit des transformations sans se demander dans quelle catégorie « artificielle » ces transformations se classent. Ces modifications peuvent porter sur l'une quelconque des caractéristiques du corps : position, mouvement, pression, température...

Cela revient à dire que la thermodynamique s'occupe de tous les phénomènes qui mettent en jeu de l'énergie, quelle que soit la forme prise par cette énergie. Le nom donné à ce chapitre de la physique rappelle simplement le rôle à part, et spécialement important, joué par cette forme particulière d'énergie qu'est la chaleur. On constate d'ailleurs ce rôle exceptionnel joué par la chaleur beaucoup plus qu'on ne l'explique théoriquement.

En effet, le deuxième principe de la thermodynamique est admis a priori. On peut évidemment penser que ce postulat est largement vérifié par ses conséquences ; il n'en reste pas moins que la chaleur garde une place spéciale. Son comportement particulier demeure

assez peu satisfaisant du fait du manque de « symétrie » de la théorie (nous n'envisageons ici la thermodynamique que sous sa forme macroscopique, puisqu'en théorie statistique on retrouve l'unité).

Le grand intérêt des raisonnements thermodynamiques provient de leur extrême généralité.

CONCEPTS DE BASE

L'un des concepts de base est celui de système. On appelle système, une portion d'univers que l'on isole (réellement ou par la pensée) du reste du monde. On peut alors distinguer « l'intérieur » et l'« extérieur » du système et définir des échanges entre les deux portions d'univers ainsi délimitées. On peut noter dès maintenant un fait remarquable : cette notion, fondamentale en thermodynamique, est également fondamentale en cybernétique.

De même, on appelle « système complètement isolé », un système isolé thermiquement et mécaniquement, c'est-à-dire qui n'échange dans un sens ou dans l'autre aucune forme d'énergie avec l'extérieur. L'analogie signalée précédemment se poursuit et l'on remarque qu'un « système isolé », au sens de la thermodynamique, est également obligatoirement un « système isolé », au point de vue de l'information. Celle-ci est, en effet, inconcevable sans un support énergétique, si petit soit-il. La réciproque est vraie (par contre, il faut remarquer qu'un système, considéré comme « pratiquement isolé », peut ne pas l'être au point de vue de l'information).

Premier principe.

Si l'on appelle,

T , le travail mécanique,

Q , la quantité de chaleur,

$T > 0$, indique que le système fournit du travail,

$T < 0$, indique que le système absorbe du travail,

$Q > 0$, indique que le système absorbe de la chaleur,

$Q < 0$, indique que le système fournit de la chaleur.

On peut énoncer le premier principe de la thermodynamique (ou principe de l'équivalence)

$$T = J Q$$

Il exprime que la quantité de chaleur fournie au système est égale — à un coefficient J près, qui est fonction des unités — au travail produit par le système (les quantités peuvent être > 0 ou < 0).

Ce principe est très général car la nature des transformations envisagées reste absolument quelconque ; le système peut, par exemple, subir aussi bien des changements d'état ou des transformations chimiques que des changements de position de ses parties.

Énergie interne.

En utilisant les mêmes notations, on peut définir la variation d'énergie interne d'un système. Lorsqu'un système passe de l'état A à l'état B, la quantité $(J Q - T)$ définit — à une constante arbitraire près — une fonction U du système, telle que l'on ait

$$U_B - U_A = J Q - T$$

Cette fonction U est l'*énergie interne* du système considéré. Si l'on applique cette définition à un système isolé, on voit que T et Q sont toujours nuls par définition. On peut par conséquent écrire que « l'énergie interne d'un système isolé est constante ». Cette énergie interne est en fait la somme de l'énergie cinétique et de l'énergie potentielle. Ce qui s'écrit (quelle que soit la forme sous laquelle est accumulée l'énergie potentielle)

$$U = C + P$$

Ces définitions entraînent l'impossibilité du mouvement perpétuel de première espèce (sous la forme de travail perpétuel). On voit immédiatement en effet que l'énergie interne ne pourrait alors que diminuer constamment.

Deuxième principe.

Le deuxième principe de la thermodynamique (ou principe de Carnot) introduit plusieurs éléments nouveaux par rapport au premier principe :

— La distinction entre « l'énergie travail » et « l'énergie chaleur ». L'un des meilleurs exemples que l'on puisse donner de cette différence est le suivant : en chauffant un ressort, on augmente son énergie interne aussi bien qu'en le comprimant et cependant ces deux opérations donnent en fait des résultats très différents.

— La distinction entre des quantités de chaleur égales, mais fournies à des températures différentes.

— L'importance théorique fondamentale des transformations réversibles. Rappelons qu'on appelle transformation réversible, une transformation dans laquelle, à chaque instant, l'équilibre est réalisé tant au point de vue thermique qu'au point de vue « mécanique ».

On admet alors le postulat suivant : *une transformation fermée (réversible ou non) d'un système qui n'échange de chaleur qu'avec un milieu, à une seule température, ne peut produire un travail positif.*

Ce qui revient à poser, a priori, l'impossibilité du mouvement perpétuel de deuxième espèce.

Les conventions faites plus haut permettent alors d'écrire que, si une transformation fermée, comprenant des échanges avec deux sources à des températures différentes, produit un travail positif, le système a reçu de la chaleur de la source chaude et en a cédé à la source froide. Ce résultat permet l'introduction de la notion de rendement, par la comparaison du résultat *utile* obtenu et ce qu'il a fallu dépenser effectivement pour le produire.

Si l'on appelle n ce rendement, on peut le définir par

$$n = \frac{T}{J Q}$$

Bien entendu cette définition ne contrevient en rien au principe de la conservation de l'énergie puisque la quantité de chaleur qui n'est pas transformée en travail est cédée à la source froide.

Le théorème de Carnot précise cette notion de rendement. On peut l'énoncer :

« Tous les cycles de Carnot fonctionnant entre les mêmes températures ont le même rendement. Cette valeur représente le maximum qu'on peut espérer obtenir, compte-tenu des températures des sources ».

Ce théorème met par ailleurs en évidence à nouveau le fait que toute irréversibilité correspond à une perte de travail mécanique. En effet, le théorème énoncé plus haut s'applique à un cycle de Carnot, c'est-à-dire un cycle réversible. Lorsque le cycle n'est pas réversible, le rendement est inférieur à celui indiqué par le théorème de Carnot.

Si on mesure alors la température sur une échelle de températures absolues, le rendement maximum s'écrit

$$n = \frac{T - T'}{T}$$

où T est la température absolue de la source chaude, et T' , celle de la source froide.

Ce théorème peut d'ailleurs s'exprimer d'une façon symétrique. Si on pose,

Q , quantité de chaleur fournie par la source à température absolue T ,

Q' , quantité de chaleur reçue par la source à température absolue T' , on peut écrire

$$\frac{Q}{T} + \frac{Q'}{T'} = 0$$

si la transformation est réversible, et

$$\frac{Q}{T} + \frac{Q'}{T'} < 0$$

si la transformation n'est pas réversible.

Ce théorème se généralise immédiatement en cas de n sources à températures absolues T_1, T_2, \dots, T_n qui fournissent, au système, des quantités de chaleur (algébriquement)

$$Q_1, Q_2, \dots, Q_n$$

Pour une transformation réversible, on aura

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} + \dots + \frac{Q_n}{T_n} = 0$$

Pour une transformation irréversible, au contraire,

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} + \dots + \frac{Q_n}{T_n} < 0$$

On peut encore généraliser, en passant à la limite, pour un nombre infini de sources et écrire

$$\int \frac{dQ}{T} = 0$$

pour une transformation réversible, et

$$\int \frac{dQ}{T} < 0$$

pour une transformation irréversible.

On peut alors, à partir de ces dernières équations, définir une grandeur S telle que, dans le cas d'une transformation réversible,

$$\int \frac{dQ}{T} = S_1 - S_0$$

S_0 étant la valeur correspondant à l'état initial et S_1 à l'état final. Cette fonction S s'appelle l'entropie du système. On voit alors aisément que $\frac{dQ}{T}$ est une différentielle exacte

$$\frac{dQ}{T} = dS$$

Dans le cas d'une transformation irréversible, on a, au contraire,

$$\int \frac{dQ}{T} < S_1 - S_0$$

Par conséquent, dans un système isolé, l'entropie ne peut qu'augmenter : elle reste constante pour les transformations réversibles et elle augmente pour les transformations irréversibles. De là, résulte qu'un système isolé ne peut pas repasser deux fois par le même état.

Cette conclusion est fondamentale et on peut même la généraliser à l'ensemble de l'univers, considéré comme un système isolé. On a, alors, les conséquences suivantes :

- l'énergie interne de l'univers est constante,
- l'entropie de l'univers croît constamment (ou en tout cas ne peut décroître à aucun moment),
- l'univers ne peut passer deux fois de suite par le même état.

ÉLÉMENTS POUR UNE GÉNÉRALISATION DE LA THÉORIE DE L'INFORMATION

OBJET

Nous avons vu précédemment que la thermodynamique s'occupe de tous les phénomènes qui mettent en jeu de l'énergie, quelle que soit la forme prise par cette énergie. De la même façon, une théorie générale de l'information traitera de tous les phénomènes mettant en jeu de l'information, indépendamment de la forme précise sous laquelle peut se représenter cette information.

Nous tenterons d'indiquer ici les principaux éléments qui doivent, selon nous, former la base d'une telle théorie générale de l'information. Les résultats obtenus présentent le même degré de généralité et d'universalité que ceux de la thermodynamique. Et c'est

précisément la généralité des règles énoncées, qui rend difficile l'évaluation de l'intérêt exact d'une telle théorie, ainsi que sa vérification.

Nous n'avons pas poussé cette étude très au-delà de l'énoncé des principes généraux, mais ces principes pourraient être appliqués aux calculs bien connus qui ont été effectués dans le cadre de la théorie désormais classique de la transmission de l'information selon Shannon. Nous rappelons d'ailleurs brièvement ci-dessous, quelques points de cette théorie.

RAPPEL DE LA THÉORIE DE SHANNON

L'essentiel des travaux de Hartley, Gabor, Ville, Wiener, Tuller, etc., sur la théorie de l'information, a porté sur la mesure et la transmission de l'information. L'œuvre de Shannon a essentiellement consisté à effectuer la synthèse de tous ces travaux, et l'essentiel de sa théorie peut se résumer par l'énoncé d'un théorème :

« La transmissibilité d'un message, dans un temps donné, est une quantité mesurable ». Cette quantité s'appelle « vitesse d'information ». Cette remarque est due à Aigrain (*op. cit.*).

Il est tout d'abord nécessaire de préciser ce que l'on entend par le terme « message ».

En effet, si l'on pouvait choisir une fonction complètement arbitraire, pour transmettre de l'information, il n'y aurait plus de problème.

On peut aisément le démontrer ⁽¹⁾.

Considérons en effet le problème suivant :

N personnes ont à transmettre N fonctions du temps $S_1(t)$, $S_2(t)$, ... $S_n(t)$.

On met à leur disposition une fonction $S(t)$ unique. Le procédé de transmission ne peut être que le suivant : $S(t)$ sera une fonctionnelle de S_1, S_2, \dots, S_n

$$S(t) = F(S_1, S_2, \dots, S_n)$$

A partir de S , à la réception, il sera nécessaire de reconstituer S_1, S_2, \dots, S_n . Pour cela, on fera apparaître N fonctionnelles de S

⁽¹⁾ Jean VILLE, *Quantité d'information ; exposé théorique*. Ex *La Cybernétique*, op. cit., pp. 151-161.

$$\begin{aligned}\sigma_1 &= G_1(S) \\ \sigma_2 &= G_2(S) \\ \sigma_n &= G_n(S)\end{aligned}$$

La transmission sera correcte, si $\sigma_1, \sigma_2, \dots, \sigma_n$ reproduisent S_1, S_2, \dots, S_n .

Il faudrait, pour cela, disposer d'une fonction complètement arbitraire. Comme une telle fonction est impossible en pratique, on se limite en fait à une classe de fonction S , en admettant que « dans un intervalle de temps T , dans des conditions de transmission données, deux fonctions $S(t)$ ne sont discernables que si elles sont suffisamment différentes ».

Il est donc nécessaire de définir la discernabilité de deux signaux, car en l'absence d'une telle définition, il est impossible d'envisager une mesure précise de l'information contenue dans un message.

Pour cela, on définit un message par

— le signal que fait correspondre le transducteur d'entrée de la voie de communication considérée au message à transmettre ;

— l'ensemble des transformations que l'on peut faire subir à ce signal, sans changer le message que lui fera correspondre le transducteur de sortie de la voie de communication considérée.

L'actuelle théorie de l'information considère deux types de transformations particulièrement importantes.

a) Les transformations de type A , ou transformation de « redondance ». Les transformations de ce type forment un groupe.

Si, parmi les transformations applicables au signal, il y en a de type A , cela signifie que le récepteur est insensible à une certaine caractéristique du signal.

b) Les transformations de type B ou d'« imprécision ». Ces transformations signifient qu'il y aura toujours une limite à la précision avec laquelle on pourra distinguer des messages correspondant à 2 signaux très peu différents l'un de l'autre.

Pour préciser ce type de transformation, on définit la notion de *distance* entre 2 signaux.

Cette distance est la valeur moyenne $d(x, y)$ d'une quantité de la forme

$$D(x, x - y)$$

où x et y sont les 2 signaux considérés ; la fonction D doit jouir des propriétés suivantes :

$$D(x, x - y) = D(y, y - x)$$

$$D \geq 0$$

$$D(x, 0) = 0$$

Une transformation de type B fera correspondre à un signal x , un signal y , tel que

$$D(x, y) \leq \epsilon$$

Dans sa forme actuelle, la théorie des communications est limitée au cas où on peut obtenir l'ensemble des transformations applicables au signal, en appliquant d'abord une transformation de type A , puis une transformation de type B . Si T est une transformation applicable au signal, on doit toujours avoir

$$T = B A$$

On peut alors tenter de calculer la quantité d'information.

CAS DES SIGNAUX DISCRETS

On appelle signal discret, un signal tel que toute portion de durée finie représente un choix entre un nombre fini de possibilités.

Soit un signal discret, de durée finie, qui représente un choix entre N_1 possibilités, toutes différentes et également probables.

Soit un autre signal discret, de durée finie, qui représente un choix entre N_2 possibilités.

Pour que l'on ait affaire à une quantité mesurable, il faut que

$$R(N) = R(N_1) + R(N_2)$$

Mais on voit par ailleurs que la juxtaposition des 2 signaux donne un signal qui représente un choix entre $N_1 \times N_2$ possibilités.

On aura donc

$$R(N_1) + R(N_2) = R(N_1 \times N_2)$$

Ce qui impose

$$R = k \log N$$

La constante k est déterminée par le choix de l'unité de mesure.

Si l'on choisit comme unité le choix entre deux possibilités (bit ou hartley), on aura

$$R = \frac{k \log N}{k \log 2} = \log_2 N$$

Si l'on considère des cas plus complexes, on peut être amené à considérer, non plus la quantité totale d'information de l'ensemble d'un signal, mais la quantité moyenne d'information transmise par unité de temps. On a alors :

$$R = \lim_{T \rightarrow \infty} \frac{1}{T} R(T)$$

$$T \rightarrow \infty$$

où $R(T)$ est la quantité d'information transmise dans le temps T .

Si, par exemple, le signal de durée T ne représente un choix entre $N(T)$ possibilités, toutes différentes et également probables à

$$R = \lim_{T \rightarrow \infty} \frac{1}{T} \log_2 N(T)$$

$$T \rightarrow \infty$$

DÉFINITIONS

Un certain nombre de concepts de base sont identiques ou parallèles à ceux qui ont été indiqués précédemment pour la thermodynamique.

NOTION DE SYSTÈME

La notion de système, dont nous avons vu l'importance en thermodynamique, est valable, sans modification, pour la théorie de l'information. Dans ce cas, en effet, on appelle également « système », une portion d'univers que l'on isole (réellement ou par la pensée) du reste du continuum. On peut alors distinguer l'intérieur et l'extérieur de ce système et définir des échanges d'information entre ces deux portions d'univers. Nous désignerons un système quelconque par ε .

Un système ε est dit « isolé », au point de vue informationnel, lorsqu'il n'échange aucune information avec l'extérieur, dans un sens ou dans l'autre.

Il est remarquable qu'un système isolé du point de vue informationnel doive l'être aussi du point de vue énergétique, la liaison entre l'information et l'énergie ne paraissant en effet pas discutable ⁽¹⁾.

(1) Une indéniable symétrie apparaît entre l'énergie et l'information. Il est possible de diminuer, presque indéfiniment, l'énergie nécessaire à la transmission d'une infor-

L'INFORMATION

Tout essai pour définir de façon précise l'« information » se heurte à des difficultés considérables. Ces difficultés sont parfaitement normales, étant donné le degré d'abstraction des termes qui doivent être utilisés. Il en est, en effet, de même, chaque fois que l'on veut s'attaquer à une notion réellement fondamentale. On se heurte au même type de difficulté lorsque l'on tente de définir les entités comme le temps, l'énergie, l'espace.

L'information doit se situer au même niveau puisqu'elle nous apparaît dès maintenant comme étant au moins aussi fondamentale que l'énergie. Cette complexité croît encore si l'on remarque que l'information ne paraît pas pouvoir être considérée, dans l'absolu, comme une notion isolée. Tout comme actuellement l'espace et le temps apparaissent indissolublement liés, l'énergie et l'information ne peuvent pas être séparés absolument. Malgré cette certitude d'ailleurs, la nature exacte de cette liaison est fort loin d'être claire.

En fait, malgré ces obstacles, on peut tenter de donner de l'information une définition approchée, ou tout au moins de la cerner.

Il est tout d'abord possible de préciser ce qu'elle n'est pas. L'information, bien que liée à l'énergie, peut, en effet, lui être opposée. On peut exprimer cette opposition en disant que, dans un flux d'énergie, l'information est représentée en quelque sorte par la partie non énergétique du phénomène (par exemple, la mesure de l'amplitude d'un flux et sa variation dans le temps).

On voit de plus, que l'information est obligatoirement liée au temps et à l'espace. On ne conçoit, en effet, pas d'information en l'absence d'un déroulement de signal dans le temps ou dans l'espace (et généralement les deux, un signal ne pouvant jamais exister instantanément si la possibilité de transmission instantanée n'est pas admise).

mation, mais en aucun cas on ne peut concevoir la disparition pure et simple de cette énergie.

De la même façon, on peut toujours diminuer considérablement l'information portée par un flux énergétique sortant d'un système, ou y pénétrant, mais on ne peut pas l'annuler.

A la limite, l'existence de ce flux d'énergie transporte au moins une information qu'il n'est pas possible de supprimer, sans supprimer du même coup l'énergie elle-même : l'existence du système qui émet ce flux. Il semble, par conséquent, qu'un échange énergétique, en l'absence de tout échange informationnel, soit aussi inconcevable que l'inverse, et qu'un système isolé, du point de vue informationnel, doive l'être obligatoirement, du point de vue énergétique.

On peut, par ailleurs, tenter de définir l'information de façon pragmatique, c'est-à-dire en se référant à son utilisation. On pourra alors dire que l'information est l'« entité non énergétique » qui est à la base de toute action, ou si l'on préfère, l'entité qui « module » l'énergie disponible à l'état brut pour la transformer en action orientée vers un but déterminé. Plus l'action à obtenir est complexe et différenciée, plus la quantité d'information mise en œuvre sera importante.

On peut encore dire, dans la même optique, que l'information est l'entité qu'un organisme (au sens généralisé) reçoit du monde extérieur, portée par un flux énergétique, mais non confondue avec lui, pour lui permettre d'adapter son action à ce monde extérieur. Dans ces conditions, plus un organisme sera apte à recevoir l'information, plus sa conduite aura de chance d'être adaptée au monde extérieur.

Par contre, si l'on éprouve de notables difficultés à définir le contenu exact de la notion d'information, on sait mesurer cette grandeur de façon simple. Nous avons défini précédemment la notion de quantité d'information et indiqué la marche à suivre pour son calcul dans quelques cas simples.

ÉMETTEUR D'INFORMATION

On appellera ainsi tout système susceptible de céder au monde extérieur (ou à un autre système) une quantité d'information $I > 0$. L'émetteur d'information jouera, dans la théorie générale de l'information, un rôle assimilable, en partie, à celui de la « source chaude » en thermodynamique.

RÉCEPTEUR D'INFORMATION

On appellera ainsi tout système susceptible de recevoir du monde extérieur une quantité d'information $I > 0$. Le récepteur d'information jouera, entre autre, le rôle dévolu à la « source froide » en thermodynamique.

QUANTITÉ D'INFORMATION INTERNE D'UN SYSTÈME

On appellera quantité d'information interne d'un système, la quantité totale d'information qui se trouve à l'intérieur de ce système, quelle que soit la forme sous laquelle cette information y est stockée.

INFORMATION UTILE

Un message apportera une quantité d'information utile, lorsqu'après réception de ce message, la quantité d'information utile du système récepteur aura une valeur supérieure à la valeur qu'elle possédait avant la réception du message.

RENDEMENT INFORMATIONNEL

On appellera rendement informationnel d'un message, la valeur

$$n = \frac{\text{Quantité d'information utile du message}}{\text{Quantité globale d'information contenue dans le message}}$$

Par définition

$$n \leq 1$$

DOMAINE D'APPLICATION
ET PRÉCISION DE L'INFORMATION

Contrairement à ce que considèrent nombre de spécialistes, il ne paraît pas possible d'élaborer une théorie générale de l'information, sans tenir compte des objets ou des domaines sur lesquels porte cette grandeur, sans « personnaliser » en quelque sorte cette information.

On voit immédiatement, en effet, qu'il ne peut exister, dans la réalité, d'information indépendante. Il s'agit toujours « d'information au sujet de quelque chose ». De façon plus précise, elle se présente même, ou peut être mise, sous la forme « réponse à une question bien déterminée ».

Deux quantités d'information données ne pourront donc être comparées, et surtout combinées, que si elles portent sur le même objet.

Il s'agit là, de faits d'expérience courante, et l'on peut prendre des exemples dans divers domaines.

— Dans un mécanisme quelconque, les informations reçues portent sur une grandeur déterminée et répondent à une question bien précise. Elles sont d'ailleurs fournies par des capteurs dont le but est d'être sensibles à un seul paramètre. C'est ainsi qu'un capteur délivrera une grandeur, fonction du paramètre que le mécanisme doit surveiller ou réguler. Ex. : pression, température, épaisseur, composition d'un mélange, etc.

— Lorsqu'un examinateur interroge un candidat, son but est de vérifier le contenu informationnel de la mémoire du candidat. Pour cela, il pose des questions précises, et ne tient compte que des réponses adéquates.

— Un journaliste envoyé par son journal, câble le récit d'un événement qui vient de se produire dans la région où il se trouve. En apparence, il ne répond pas à une question. En fait, tout comme le capteur du servomécanisme, il répond au programme qui lui a été fixé. Dans son cas, la question posée est : « Qu'y a-t-il de nouveau ? ».

Il envoie un message dès que cette question comporte une réponse positive. Dans son cas, le domaine est très vaste. Il se restreint déjà, si l'envoyé est un spécialiste du sport, de la politique, de l'actualité scientifique. Les messages gagneront alors en précision ce qu'ils pourront perdre en extension.

C'est seulement lorsque l'on s'intéresse uniquement à la transmission de l'information, que l'on peut faire abstraction de cet aspect. La quantité d'information à transmettre est alors considérée comme une donnée et le problème se réduit à obtenir à la réception, de la façon la plus rapide et la plus économique possible, un message aussi identique que possible au message de départ (compte-tenu de la définition de cette identité).

On retombe alors sur la théorie de Shannon qui représente finalement le « point de vue du télégraphiste », c'est-à-dire, s'intéresse au volume d'information à transmettre, indépendamment de la nouveauté et de l'intérêt que peut présenter cette information pour le destinataire.

Si l'on considère, au contraire, ce problème, d'un point de vue le plus général possible, on voit que deux quantités d'information qui ne portent pas sur le même domaine, ne pourront pas normalement être additionnées, ni même comparées entre elles.

Il s'agit là d'un problème identique à celui qui se pose lorsque l'on applique à des cas pratiques les règles de l'arithmétique. On ne peut en effet additionner des quantités que lorsqu'elles portent sur des objets identiques.

L'analogie peut d'ailleurs être développée. Il suffit pour cela, de remarquer que la théorie de Shannon — définissant les lois de combinaison de l'information — se conduit exactement comme l'arithmétique qui combine des nombres sans se soucier de ce que pourraient éventuellement représenter ces nombres.

Au contraire, les lois de combinaison que nous énonçons plus loin, jouent le même rôle que la numération appliquée à des objets bien définis.

Il est toujours possible de combiner des quantités d'information en vue de leur transmission, tant que l'on ne cherche pas à connaître les objets sur lesquels porte cette information, de la même façon qu'il est toujours possible d'additionner des nombres purs. Mais il n'est, en général, plus possible de procéder ainsi lorsque l'on examine un cas « concret », en se plaçant du point de vue de l'utilisateur.

Nous retrouvons, à nouveau, le parallèle avec la thermodynamique. Si l'on se place du point de vue du premier principe, les calories peuvent s'additionner ou se soustraire de la façon la plus simple. Au contraire, après l'énoncé du second principe, les choses deviennent beaucoup moins simples. Les notions de transformation et de rendement font alors intervenir le « point de vue de l'utilisateur » et il est nécessaire de préciser la forme et les caractéristiques exactes de la quantité d'énergie étudiée.

NOTION DE SYSTÈME INFORMATIONNEL ÉLÉMENTAIRE

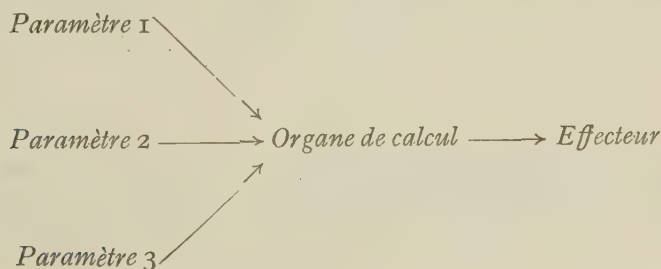
Il est nécessaire de préciser cette notion.

Considérons un système ε susceptible de recevoir de l'information venant du milieu extérieur, ou d'autres systèmes. Cette information entre dans le système par l'intermédiaire de capteurs, définissant chacun un domaine informationnel. On divisera alors l'ensemble du domaine dans lequel ε peut recevoir de l'information en domaines élémentaires D_1, D_2, \dots, D_n (chaque domaine ne pouvant contenir que des quantités d'information qui concernent une seule question).

On peut également considérer le système comme formé de systèmes élémentaires du point de vue informationnel $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ (chacun d'eux ne possédant qu'un seul capteur).

Il est nécessaire de remarquer que la décomposition en système élémentaire, n'obéit pas forcément aux mêmes lois qu'une décomposition « matérielle ». En particulier, un même élément du système peut parfaitement figurer dans plusieurs systèmes élémentaires.

Considérons à titre d'exemple, le schéma ci-dessous



Il peut se décomposer en 3 schémas du type suivant

Paramètre \longrightarrow *Organe de calcul* \longrightarrow *Effecteur*

C'est finalement à un système élémentaire de ce type que s'adresse un message donné, et c'est à l'intérieur de ces systèmes que s'appliquent les formules de combinaison et de rendement indiquées plus loin, pour la quantité d'information.

NOTION DE PRÉCISION D'INFORMATION

On est également amené à définir une grandeur qui jouera, vis-à-vis de l'information, le rôle joué par la température dans le cas de l'énergie calorifique. Cette grandeur pourra être appelée la « précision d'information ». Elle caractérise la quantité d'information existant sur un point donné, c'est-à-dire, le fait que la réponse à une question peut être plus ou moins complète.

Par exemple, si l'on cherche à transmettre la valeur du nombre π , la réponse $\pi = 3,14$ apporte une quantité d'information moindre sur cette valeur (et par conséquent possède un degré de précision moindre) que la réponse $\pi = 3,14\ 15\ 92$. Il est essentiel de remarquer que cette notion de degré de précision ne pourra s'appliquer qu'à une information bien déterminée ou à un domaine strictement limité, c'est-à-dire, à l'intérieur d'un système élémentaire du point de vue informationnel tel que nous venons de le définir.

Dans le cadre de l'analogie avec la température, on considérera donc en général un système complexe ε comme équivalent, non à une source unique à température T , mais plutôt à un système contenant n sources distinctes à des températures T_1, T_2, \dots, T_n .

On peut pousser plus loin l'analogie avec la chaleur et la température. On sait, en effet, qu'une source ne peut, du moins si on se place d'un point de vue statistique, recevoir une énergie > 0 , que d'une source à une température supérieure à la sienne.

On constate, de la même façon, qu'une « source d'information », c'est-à-dire un système isolé contenant de l'information et susceptible d'en recevoir (et éventuellement d'en émettre), ne pourra recevoir une information supérieure à 0, que d'une source possédant un niveau de précision supérieur au sien.

LOIS DE COMBINAISON DE L'INFORMATION

Si l'on considère une ligne de transmission (sous sa forme la plus générale), dans laquelle est envoyée de l'information, il ne paraîtra pas, en général, nécessaire de doter cette ligne d'une mémoire qui lui permette d'enregistrer l'ensemble des messages qui y circuleront.

Cette restriction, parfaitement normale étant données les caractéristiques des systèmes de transmission réels, interdit en fait à la théorie de la transmission de l'information d'énoncer les lois de combinaison réelles de l'information. C'est justement l'obligation pour la théorie de Shannon de postuler implicitement cette mémoire (avec l'hypothèse des retards de transmission illimités) qui est l'un des points faibles de cette théorie.

En effet, en l'absence de mémoire, on ne peut que considérer les messages qui passent, en un point donné, au cours du temps, et la quantité globale d'information qui parcourt la ligne, durant une période T , sera égale à

$$\int_0^T dI$$

dI représentant la quantité d'information qui passe au point P entre les instants t et $t = dt$ (cette valeur étant généralement une fonction de t). On simplifie habituellement la formule précédente, en considérant la valeur moyenne par unité de temps de la quantité d'information. On arrive alors à l'expression de la quantité d'information transmise en moyenne, de la forme

$$R = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dI$$

dont le calcul est généralement plus simple que celui de $\int_0^T dI$.

Si, au contraire, on considère un système ϵ capable de recevoir ou d'émettre de l'information, et capable d'utiliser cette information, on doit obligatoirement le supposer doté d'une « mémoire », si rudimentaire ou si implicite soit-elle.

Considérons alors un système ϵ (contenant N domaines élémentaires D_1, D_2, \dots, D_n , avec des quantités d'information I_1, I_2, \dots, I_n , de précision P_1, P_2, \dots, P_n). Ce système reçoit de l'extérieur une information I_x de précision P_x (portant sur le domaine D_i par exemple). Ce domaine D_i contient déjà une quantité d'information I_i à la précision P_i . On peut alors écrire que

$$(I_i + I_x) = I_i \text{ pour } I_x \leq I_i$$

$$(I_i + I_x) = I_x \text{ pour } I_x \geq I_i$$

en exprimant par la parenthèse l'opération d'addition de 2 quantités d'information « personnalisée » comme nous avons vu précédemment.

On remarquera immédiatement, comme conséquences de ces lois appliquées à des cas particuliers évidents, que

— Si le système ϵ ne contient rien à l'instant 0, l'ensemble du message sera neuf et l'on aura alors

$$I_{\text{final}} = I_x$$

— Si le système ϵ contient à l'instant 0, l'ensemble de l'information contenue dans le message, celui-ci n'est pas neuf, et l'on retrouve bien

$$I_x + I_x = I_x$$

c'est-à-dire que le message n'a rien apporté.

— Si l'on considère un système sans mémoire (c'est-à-dire une ligne de transmission), on voit que I initial sera toujours nul, et par conséquent le message toujours neuf. On a donc alors, si l'on considère l'information qui entre dans le système au cours du temps

$$I_{\text{total}} = \epsilon I_n$$

le ϵ désignant la somme arithmétique. On retrouve donc bien la théorie de Shannon.

Remarque. En théorie macroscopique de l'information, la soustraction de deux quantités d'information est une notion normalement dépourvue de sens ; aussi, ne la considérerons-nous pas ici.

PREMIER PRINCIPE DE LA THÉORIE DE L'INFORMATION

Le premier principe de la thermodynamique (ou principe de l'équivalence) n'est en fait que l'affirmation de l'impossibilité du mouvement perpétuel de première espèce.

Il se ramène en effet à énoncer que, quelle que soit la forme prise par une quantité d'énergie donnée, cette quantité d'énergie ne peut jamais augmenter. Il en découle évidemment l'impossibilité de produire « gratuitement » de l'énergie sous quelle que forme que ce soit.

On pourra, parallèlement, appeler 1^{er} principe de la théorie de l'information, l'énoncé de l'impossibilité de l'augmentation de la quantité d'information interne d'un système isolé.

C'est cette production d'information nouvelle, sans contrepartie, que certains auteurs ont appelée « mouvement perpétuel de 3^e

espèce » et ils ont fait grief à la théorie de Shannon de ne pas démontrer l'impossibilité de cette variante du mouvement perpétuel.

En fait cette impossibilité découle immédiatement des lois de combinaison que nous avons énoncées. Soient en effet deux systèmes (ε_1 et ε_2) contenant respectivement des quantités d'information I_1 et I_2 appliquées à des domaines D_1 et D_2 .

On a les notations suivantes :

I_1 = quantité d'information contenue dans le système ε_1 . Par hypothèse $I_1 \geq 0$,

I_2 = quantité d'information contenue dans ε_2 . De même $I_2 \geq 0$ par hypothèse.

Admettons alors que ε_2 envoie la quantité d'information I_2 qu'il contient au système ε_1 . Si $I_2 > I_1$ (avec les restrictions que nous avons indiquées précédemment au sujet des domaines d'application), il en résulte pour ε_1 un accroissement réel d'information ΔI tel que l'on pourra écrire, en appliquant les lois de combinaison, avec I_1 et $I_2 \geq 0$

$$\Delta I = I_2 - I_1$$

(en admettant ici $D_1 = D_2$)

Cette égalité implique que ΔI n'existe que pour $I_2 > I_1$ puisque $\Delta I < 0$ est une impossibilité, un système ne pouvant, en général, perdre d'information par une opération de ce type.

Par ailleurs, on voit immédiatement que cette valeur ΔI dépend uniquement de l'état initial et de l'état final, et non de la façon dont est transmise la quantité d'information I_2 . Cette caractéristique découle des lois de combinaison et en particulier de

$$(I_i + I_j) = I_i$$

On peut alors écrire

$$\Delta I = I_2 - I_1 \text{ si } I_2 > I_1$$

$$\Delta I = 0 \quad \text{si } I_2 \leq I_1$$

En appelant

$(I_1)_0$, l'information interne de ε_1 dans l'état initial,

$(I_1)_1$, l'information interne de ε_1 dans l'état final et en remarquant que l'on a obligatoirement

$$(I_1)_1 \geq (I_1)_0$$

(en rappelant que l'on ne tient pas compte de l'entropie, cette restriction ayant été justifiée par ailleurs), on a

$$\begin{aligned}
 (I_1)_1 - (I_1)_0 &= \Delta I \\
 &= I_2 - (I_1)_0 \text{ si } I_2 > I_1 \\
 &= 0 \quad \quad \quad \text{si } I_2 \leq I_1
 \end{aligned}$$

La démonstration de l'impossibilité du mouvement perpétuel de 3^e espèce est alors évidente. On a en effet, par hypothèse,

$$I_2 = 0$$

puisque l'on a affaire au seul système ε_1 qui n'emprunte aucune information au milieu extérieur.

Comme par ailleurs :

$$I_1 \geq 0$$

on a obligatoirement

$$I_2 \leq I_1$$

On est donc placé dans le 2^e cas des formules précédentes et

$$\Delta I = 0$$

Il ne peut donc y avoir aucune augmentation de l'information interne du système ε_1 en l'absence d'apport d'information du milieu extérieur.

L'intérêt de cette démonstration (comme celle de l'impossibilité du mouvement perpétuel en thermodynamique) réside dans sa très grande généralité. Étant donné, en effet, que les hypothèses de départ n'entraînent aucune restriction, cette démonstration peut s'appliquer à tout système, si complexe soit-il (qu'il s'agisse par exemple d'une machine, du cerveau humain ou de l'univers dans son ensemble).

On voit, d'autre part, que cette démonstration est une conséquence évidente des lois de combinaison de l'information.

L'ensemble de la théorie repose donc finalement sur ces lois et n'en est que le développement.

PARALLÈLE ENTRE INFORMATION CHALEUR ET TRAVAIL

LOI DE LA CONSERVATION DE LA QUANTITÉ D'INFORMATION

Si nous poursuivons le parallèle avec la thermodynamique, on voit que nous pouvons comparer l'« information utile » et le travail, tandis que l'information en général (celle contenue dans les messages par exemple) est comparable à l'énergie calorifique.

Cette assimilation implique immédiatement la possibilité du passage de l'une à l'autre forme et la notion de rendement.

Ce rendement, tout comme en thermodynamique, dépendra des caractéristiques des systèmes en présence.

Dans un système ϵ contenant une quantité d'information I , on peut, dans certains cas, dire que toute cette information est utilisable. En effet, cette utilisation dépend des caractéristiques du système, ainsi que de celles du milieu extérieur (ou du système extérieur) sur lequel le système doit agir.

Considérons un message arrivant de l'extérieur et pénétrant dans le système ϵ . On dira qu'il apporte une certaine quantité d'information utile si l'arrivée de ce message augmente la quantité d'information I contenue primitivement dans le système ϵ .

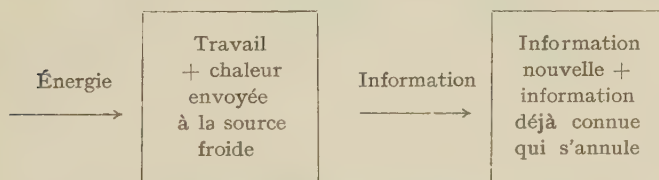
Sur ce point, se manifeste une différence fondamentale entre le comportement de l'énergie et celui de l'information.

— Dans le cas de l'énergie, les quantités d'énergie s'ajoutent algébriquement, tout au moins, lorsqu'on considère uniquement l'énergie interne du système, et non son énergie utilisable.

— Dans le cas de l'information, seule la partie « nouvelle » de l'information reçue s'ajoute, et (dans l'hypothèse macroscopique) l'information émise ne se retranche jamais.

Il s'agit là d'un fait essentiel puisque les lois de combinaison et celles du rendement se déduisent immédiatement de cette remarque.

On peut traduire sous la forme du tableau suivant



Il y a une difficulté apparente dans le fait que l'on semble pouvoir produire autant d'information à transmettre qu'on le désire, ou plus exactement, qu'un système, contenant une quantité d'information I donnée, peut, en général (et à condition d'être conçu en conséquence), émettre tout ou partie de cette quantité d'information sans diminuer en rien sa quantité d'information interne. Il pourra même généralement l'émettre autant de fois qu'on le désirera, simplement en dépensant de l'énergie en conséquence.

Ces faits semblent prouver que l'information n'obéit pas au principe de conservation. En fait, il n'en est rien, ce principe revêt simplement une forme un peu différente de celle à laquelle nous

sommes habitués pour l'énergie. On peut le voir en appliquant à ce problème les lois de combinaison de l'information.

Pour cela, considérons deux systèmes 1 et 2. On peut, sans restreindre la généralité du raisonnement, faire les hypothèses suivantes sur le contenu informationnel des deux systèmes :

ε_1 contient une quantité d'information I_1 ,
 ε_2 est vide ($I_2 = 0$).

Supposons alors que ε_1 envoie à ε_2 des messages (sous une forme telle, que ε_2 puisse les recevoir) contenant toute l'information I_1 du 1^{er} système.

Après l'opération, le système 2 contient une quantité d'information

$$I_2 = I_1$$

Il semble donc ne pas y avoir conservation de la quantité d'information. Il n'en est rien. Si, en effet, nous englobons ε_1 et ε_2 dans un système unique ε_3 , ce dernier système contient une quantité d'information mais, $I_2 = I_1$, d'où

$$\begin{aligned} I_3 &= (I_1 + I_2) \\ I_3 &= (I_1 + I_1) = I_1 \end{aligned}$$

c'est-à-dire la quantité d'information contenue d'abord dans ε_1 .

On voit donc, que les lois de combinaison permettent l'extension à l'information du principe de conservation.

La pseudo-difficulté que nous avons soulevée ne vient finalement que de l'application illégitime à l'information des lois de combinaison valables pour l'énergie.

SECOND PRINCIPE DE LA THÉORIE DE L'INFORMATION

On peut également énoncer pour l'information un théorème identique à celui qui exprime le second principe de la thermodynamique pour l'énergie.

Dans le cas de l'information, ce théorème s'introduit de façon beaucoup plus aisée qu'en thermodynamique.

Il est essentiel, tout d'abord, de bien mettre en évidence la différence existant entre deux quantités d'information égales, mais émises à des degrés de précision différents. Ce qui est d'ailleurs impossible, si elles ont la même extension en domaine d'application.

En effet, tout comme le 1^{er} principe de la thermodynamique ne fait aucune différence entre les calories, quelle que soit la température de la source qui les fournit, la théorie de la transmission de l'information ne s'intéresse qu'au nombre de hartleys transmis et pas à leur niveau de précision (ou à leur domaine d'application).

Or, l'examen des faits de l'expérience courante permet d'affirmer que, tout comme les calories ont une valeur d'autant plus grande qu'elles sont fournies à une température plus élevée, les hartleys ont une valeur « potentielle » d'information d'autant plus élevée qu'ils sont fournis à un niveau de précision plus grand. Cette considération introduit automatiquement la notion de rendement informationnel (parallèle à la notion de rendement énergétique).

On peut mettre en évidence cette différence sur des exemples pratiques simples.

Pour faire comprendre la différence entre des calories à des températures différentes, on trouve dans un traité de thermodynamique des plus classiques, cet exemple simple, mais « parlant » :

« Une source d'eau tiède peut fournir, à des corps pris à la » température ambiante, un nombre énorme de calories ; toutes » ces calories ne peuvent faire cuire un œuf, ce que l'on fait avec » un nombre bien moindre de calories fournies à température » plus élevée ».

Il en est de même en information. Supposons que l'on désire régler une horloge à une seconde près.

On dispose pour cela de 2 horloges étalons qui envoient des signaux horaires :

- la 1^{re} assure une précision de ± 10 secondes,
- la 2^{me} assure une précision de $\pm 0,1$ seconde.

Il est possible de recevoir un volume énorme d'information de la 1^{re} sans pouvoir jamais régler l'horloge à 1 seconde près, alors qu'un seul message de la 2^{me}, judicieusement utilisé, peut permettre d'effectuer ce réglage.

On peut donc étendre le raisonnement utilisé en thermodynamique et dire que les hartleys issus de la 1^{re} source ont un rendement = 0, tandis que ceux issus de la seconde ont un rendement > 0 ⁽¹⁾.

(1) La théorie de l'information se présente d'ailleurs à ce point de vue comme beaucoup plus simple que celle de l'énergie.

Considérons, en effet, l'exemple précédent. Le système contenant l'horloge à

On peut prendre un autre exemple dans un domaine différent : celui de la transmission des connaissances entre les hommes.

Supposons un professeur qui désire enseigner à un auditoire, la valeur du nombre π . Il peut, par exemple, faire un historique rapide et inscrire au fur et à mesure au tableau noir, les chiffres admis aux différentes époques.

Le chiffre 3 n'apportera en fait, que fort peu d'information utile, peu de gens l'ignorant ; par contre, si l'on considère le nombre

$$3,14\ 15\ 92\ 65\ 35\ 89\ 7\ldots$$

le chiffre 7 apportera une information importante, peu de gens le connaissant en fait (son utilité réelle est un autre problème). Là, encore, les hartleys du 1^{er} cas auront un rendement très faible, ceux du 2^e cas, un rendement fort.

Cependant, du point de vue de l'actuelle théorie de l'information, si l'on considère uniquement le nombre de hartleys nécessaires à la transmission, on arrive à la même valeur dans les deux cas.

FORMULE DU RENDEMENT INFORMATIONNEL

Ces considérations peuvent se traduire de façon précise, par l'énoncé pour l'information, d'un théorème équivalent au théorème de Carnot, régissant le rendement des échanges énergétiques.

Considérons deux systèmes ε_1 et ε_2 .

régler reçoit au temps T_1 une quantité d'information I_1 émanant de l'horloge dont la précision est de ± 10 secondes. Cette quantité d'information s'ajoute à ce que l'on connaît sur le fonctionnement de l'horloge.

Au temps T_2 le système reçoit une quantité d'information I_2 envoyée par l'horloge de précision ± 10 secondes. Cette quantité d'information n'est pas « neuve » puisqu'elle pouvait être prévue.

On a bien alors

$$I_1 + I_2 = I_1$$

En fait, compte tenu de la forme des lois de combinaison, on ne peut même pas parler réellement de quantité d'information plus grande à degré de précision inférieure, une telle conception étant contradictoire étendue au même domaine.

On a en effet obligatoirement

$$\begin{aligned} I_1 &= I_2 = \dots = I_n \\ I_1 + I_2 + \dots &= I_1 \end{aligned}$$

Donc, la quantité totale d'information émise par l'horloge n'est finalement que I_1 .



Il s'agit de systèmes isolés — au sens thermodynamique aussi bien qu'au sens informationnel. Si maintenant on couple les deux

systèmes, on constate que

— ϵ_1 est capable d'émettre (d'une façon et sous une forme quelconque) ou du moins, contient un volume déterminé d'information. Cette quantité d'information représente l'ensemble des informations contenues, de façon disponible, dans le système, éventuellement après tous calculs, raisonnements, opérations nécessaires, sans aucun apport de lois ou de calculs extérieurs (il s'agirait alors d'un apport d'information) ;

— il en est de même de ϵ_2 .

Dans ces conditions, on dira que le système ϵ_1 a reçu une certaine quantité d'information utile, de la part du système ϵ_2 , si, après réception des messages issus de ϵ_2 , il est susceptible d'émettre à son tour, une quantité d'information supérieure à celle qu'il aurait pu émettre avant l'échange.

Cette quantité d'information n'est pas forcément égale à celle mesurée par l'actuelle théorie qui se borne à considérer l'information du point de vue de sa transmission. Elle ne peut que lui être \leq .

Elle doit même lui être $<$ dans le cas général.

Le rendement informationnel des messages reçus par ϵ_1 pourra alors être défini comme le rapport de la quantité d'information utile à la quantité d'information apportée par les messages.

$$\eta = \frac{\text{information utile}}{\text{information totale}} \leq 1$$

On voit que ce rendement est fonction de la quantité d'information déjà contenue dans le système ϵ_1 et que l'on a défini précédemment comme étant l'information interne du système ϵ_1 .

Finalement, compte tenu des notations

$$\eta = \frac{|I_2| - |I_1|}{I_2}$$

Remarque : cette formule s'écrit en réalité

$$\eta \leq \frac{|I_2| - |I_1|}{I_2}$$

c'est-à-dire qu'elle indique seulement la valeur maximum que le rendement ne peut en aucun cas dépasser pour deux systèmes

donnés. Elle ne permet nullement d'affirmer que le rendement sera effectivement égal à cette valeur dans tous les cas. Il lui sera même généralement inférieur dans les cas envisagés pratiquement. Le parallèle avec le théorème de Carnot est donc complet.

On peut prendre des exemples pratiques dans des domaines divers pour illustrer cette notion de rendement informationnel. L'un des plus nets est celui de la presse qui est considérée par le langage courant, comme étant le plus typique.

La rédaction d'un quotidien reçoit généralement des dépêches provenant de plusieurs agences ou sources diverses. Elle peut ainsi recevoir plusieurs fois la même nouvelle. Du point de vue de la « théorie de la transmission de l'information » (théorème de Shannon), et du point de vue de l'exploitant du réseau de transmissions, ces dépêches (que l'on suppose exactement identiques pour plus de simplicité) transporteront des quantités d'information égales et qui s'ajoutent.

En fait, si l'on s'en tient à la définition que nous venons d'énoncer, on voit que la 1^{re} dépêche qui « tombera » dans la salle de rédaction apportera de l'information, alors que les autres n'en apporteront pas, puisqu'elles ne permettront plus d'augmenter la quantité d'information qui « ressortira » de la salle de rédaction, c'est-à-dire qui sera imprimée. Dans ce cas le raisonnement courant est en accord complet avec la notion de rendement informationnel que nous avons énoncée ci-dessus, et non avec le théorème de Shannon, car les messages successifs n'intéresseront personne.

On voit de même, que, dans ce cas aussi, le rendement est fonction de l'information déjà contenue dans le système « récepteur ». Les dépêches neuves elles-mêmes ne le sont généralement qu'en partie ou ne le sont que par rapport au journal ou au rédacteur. On prend donc effectivement, comme niveau de référence, l'ensemble des informations que représente la sommation de ce qui est passé dans les différents numéros du journal ou l'ensemble de la mémoire de l'équipe de rédaction.

Il est remarquable qu'en utilisant ce système de mesure de l'information on puisse chiffrer la *nouveauté* d'un message pour un système récepteur considéré, mais non son *exactitude*.

En l'absence de critères supplémentaires une affirmation fausse contient même généralement alors plus d'information apparente qu'une vraie.

PRÉCISIONS SUR LA NOTION DE RENDEMENT ET CONSÉQUENCES

Nous préciserons le résultat précédent en appliquant la notion de rendement informationnel à des cas particulièrement typiques. Certaines des conséquences ainsi mises en lumière, présentent un intérêt très général et montrent que la généralité de cette notion la place au même niveau que celles de rendement énergétique ou d'entropie.

a) Considérons un ensemble de systèmes tels que l'on ne possède à l'origine aucune information sur le message qui va être transmis ; l'ensemble du message sera alors essentiellement neuf. Dans ces conditions, à la redondance du message lui-même près, le rendement sera maximum (c'est-à-dire égal à 1).

Cela signifie que si l'on décrit un événement nouveau — dans un domaine dont on ne connaît aucune loi, ni l'état d'aucun paramètre — on pourra (à condition d'étudier suffisamment le message pour en éliminer toute redondance) atteindre un rendement de 100 %.

Il s'agit là d'un cas limite, vraisemblablement irréalisable en pratique.

b) Si, au contraire, l'événement peut être en partie prévu (c'est-à-dire si l'on peut fixer certaines limites pour les valeurs des paramètres définissant cet événement) à cause d'une connaissance au moins fragmentaire des lois qui le régissent et des « états initiaux », l'information réellement neuve du message diminuera. Cette diminution sera fonction de l'accroissement de la connaissance.

Un exemple classique permet d'illustrer cette thèse : la localisation d'un repère qui se trouve sur un segment AB .

La précision de cette localisation est limitée par la finesse des capteurs (œil, microscope, ...) et par la possibilité de transmettre les caractéristiques ainsi relevées. Ces limitations conduisent à diviser le segment AB en n zones distinctes qui représentent le maximum de zones qu'il est possible de distinguer compte tenu des moyens employés. Si l'on utilise alors, par exemple, le langage binaire pour transmettre cette information, on peut procéder de la façon suivante : désigner par convention, par le signe 0, la moitié gauche de segment considéré, et par 1, la moitié droite. En procédant par éliminations successives, on obtient alors un nombre qui indique la position du repère sur le segment, par exemple : 0 01 01 01 01...

La longueur de ce nombre est fonction de n .

Cette méthode est valable lorsqu'on n'a aucune précision sur la position du repère sur le segment AB . S'il est possible, au contraire, d'affirmer que, étant donné les lois auxquelles est soumise la position de ce repère, celui-ci doit se trouver entre n_1 et n_2 , on peut réduire l'intervalle. L'exploration utile ne portera donc plus sur l'ensemble du segment AB , mais sur $m < n$ zones distinctes.

Si l'on utilise alors le même nombre que précédemment pour positionner le repère, une partie des chiffres qui le composent sera inutile puisqu'ils ne feront que localiser le repère entre n_1 , et n_2 , ce qui est déjà connu. Ces chiffres auront donc un rendement informationnel nul, et on aurait pu éviter leur transmission en aménageant le système d'exploration de façon à limiter la recherche à la gamme que l'on savait être intéressante.

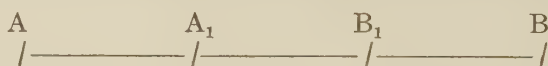
On peut également considérer ce problème sous un autre aspect.

En l'absence de toute indication permettant de préciser la localisation du repère, la probabilité pour une section quelconque de contenir ce repère est de

$$p = 1/n$$

L'information apportée par la détermination de la localisation exacte porte à 1 cette probabilité pour la section en cause et l'annule pour les autres (cette annulation ne représente d'ailleurs pas une information supplémentaire car elle découle d'un principe d'exclusion connu d'avance).

Si, au contraire, une information préliminaire permet de localiser ce repère à l'intérieur de $m < n$ sections (entre A_1 et B_1 par exemple).



la probabilité pour chacune de ces sections se trouve portée à

$$p = 1/m ; 1/m \text{ étant plus grand que } 1/n$$

Cette probabilité sera portée à 1 par la localisation exacte pour la section en cause et à 0 pour les autres.

On voit que, dans le premier cas, le rendement informationnel sera au maximum de

$$\eta = \frac{n-1}{n}$$

et dans le second cas, de

$$\eta = \frac{m-1}{m}$$

Appliquons ces formules à des cas particuliers très simples.

— Si l'on envisage le cas limite où l'on sait exactement où est situé le repère, on aura $m = 1$, d'où

$$\eta = \frac{1 - 1}{1} = 0$$

Ce résultat est évident a priori ; le message ne peut apporter aucune nouveauté, le résultat exact étant déjà connu.

— Si, au contraire, on ignore absolument tout de la position du repère, on aura

$$\eta = \frac{n - 1}{n}$$

Cette expression tend vers 1 quand $n \rightarrow \infty$. Et ceci peut se produire dans deux cas : lorsque la zone d'incertitude est étendue à l'univers entier supposé ∞ , ou bien lorsque la finesse de division possible est supposée infinie.

Si on se limite à l'univers réel, on voit que le rendement ne peut que tendre vers 1 sans jamais l'atteindre, les dimensions de l'univers étant finies ainsi que sa divisibilité.

c) On peut également appliquer le théorème du rendement à un cas particulier : celui du chaos thermique.

Supposons un système ϵ isolé dans un univers en état de chaos thermique, c'est-à-dire ayant une entropie maximum. Du point de vue macroscopique, les capteurs du système ne pourront recueillir aucun renseignement nouveau, et on aura donc de façon permanente

$$\eta = 0$$

Ce résultat, évident, est en parfait accord avec la théorie puisque cet état d'entropie maximum correspond à la probabilité maximum et au désordre maximum.

Étant à un niveau d'information \leq à celui de ϵ , puisque le plus faible possible, il ne peut évidemment lui transmettre aucune information utile (sauf son existence à l'état de chaos).

IMPOSSIBILITÉ DU MOUVEMENT PERPÉTUEL DE 4^e ESPÈCE

Tout comme la thermodynamique postule l'impossibilité de produire un travail positif à l'aide d'une seule source, quelle que soit sa température, la théorie de l'information conclut à l'impossibilité d'obtenir un rendement informationnel positif d'un message quel-

conque lorsque le système émetteur et le système récepteur présentent le même degré de précision.

La démonstration en est évidente. On a alors en effet, dans un domaine D déterminé

$$I_2 = I_1$$

donc, dans tous les cas

$$\eta = \frac{|I_2| - |I_1|}{I_2} = 0$$

GAIN INFORMATIONNEL D'UNE STRUCTURE

Il existe une objection, à première vue très importante, contre la définition actuelle de la quantité d'information d'un message.

Il s'agit du fait que la théorie de Shannon attribue dans tous les cas, la même valeur informationnelle au même message et ce, quels que soient ses destinataires et ses conditions d'émission.

Une telle conclusion est en désaccord formel avec le « sens commun ».

Il est en effet évident, par exemple, que la valeur d'information réelle d'un oui ou d'un non — 1 unité d'information — ne paraît pas être la même dans tous les cas. Il peut en effet, aussi bien s'agir d'une réponse machinale à une question sans intérêt, que du signal d'un événement historique de première grandeur.

Cette difficulté n'est finalement qu'apparente.

On peut mettre ce fait en évidence en examinant un cas particulièrement net : la valeur informationnelle du résultat numérique d'une expérience destinée à départager deux théories physiques.

Plaçons-nous, par exemple, dans le cas d'une expérience destinée à trancher entre la mécanique classique et la mécanique relativiste. Aux faibles vitesses, les résultats calculés par les deux méthodes pour une expérience quelconque, sont très voisins, c'est-à-dire qu'ils sont exprimés par deux grandeurs ayant une partie en commun.

Soit la grandeur x qui devrait prendre, en mécanique classique, la valeur

$$x_c = 1,23\ 25\ 43$$

et en mécanique relativiste, la valeur

$$x_r = 1,23\ 25\ 49$$

(les chiffres choisis sont purement imaginaires).

Tant que la précision des mesures est insuffisante, le choix entre les deux théories n'est pas possible.

Lorsque, au contraire, la précision sera devenue suffisante, la connaissance exacte du dernier chiffre permettra le choix entre deux théories dont chacune représente les quantités très importantes d'information.

Or, si l'on s'en réfère à la définition de la quantité d'information, le chiffre en litige ne peut transmettre plus d'information que les chiffres précédents, qui n'apportaient en fait aucune information permettant de trancher le débat, ou qu'un chiffre décimal arbitraire.

En fait, malgré les apparences, il n'y a là aucun paradoxe. Ce processus est exactement comparable à celui qui se déroule en thermodynamique lors d'une explosion.

Dans ce cas, en effet, l'énergie de l'étincelle (extrêmement minime) déclenche la mise à feu d'un explosif, et libère ainsi une énergie très importante par rapport à l'énergie de commande. Mais l'énergie de l'explosif a été accumulée précédemment au cours d'opérations qui étaient toutes endothermiques, et le rendement final, inférieur à 1, suit toujours le principe de Carnot.

De la même façon ici, la transmission finale d'information « libère » une importante quantité d'information (accumulée au cours d'opérations vérifiant toutes la loi du rendement) et cette loi du rendement n'est, par conséquent, en échec à aucun moment.

On peut encore comparer à ce qui se passe pour l'énergie dans le cas d'une valve ou d'un amplificateur : un signal d'entrée à faible énergie commande une énergie importante à la sortie.

On doit alors parler, non plus de rendement informationnel, mais bien de coefficient de multiplication pour un schéma informationnel déterminé, tout comme on parle de gain ou de coefficient de multiplication dans le cas d'un circuit énergétique.

REMARQUE SUR LA NOTION DE QUANTITÉ D'INFORMATION

On peut montrer que les lois de combinaison de l'information et la formule du rendement informationnel ne sont pas en contradiction avec la définition mathématique de la quantité d'information qui est l'un des points les mieux établis et les plus généraux de la théorie de Shannon.

Cette définition s'intègre au contraire naturellement dans les éléments que nous venons d'exposer.

Plaçons-nous dans le cas d'un signal discret de durée finie. Un tel signal représente un choix entre N possibilités. Un résultat classique indique que la quantité d'information R est alors donnée par la formule

$$R = k \log N$$

où k est une constante qui ne dépend que des unités de mesure choisies.

Soit alors un signal qui peut se décomposer en signaux élémentaires du type précédent représentant chacun un choix N_1, N_2, \dots, N_p possibilités (quelle que soit par ailleurs la méthode de décomposition utilisée).

On pourra écrire que le signal global représente un choix entre N possibilités, tel que

$$N = N_1 \times N_2 \times \dots \times N_p$$

Par application de la formule précédente, on aura pour la quantité totale d'information R de ce signal

$$\begin{aligned} R &= k \log N \\ &= k \log (N_1 \times N_2 \times \dots \times N_p) \\ &= k (\log N_1 + \log N_2 + \dots + \log N_p) \\ &= R_1 + R_2 + \dots + R_p \end{aligned}$$

Cette formule est valable, dans tous les cas, dans la théorie de Shannon, puisque, dans ce cas, les lois de combinaison se ramènent à

$$I = (I_1 + I_2) = I_1 + I_2$$

Si nous ramenons le raisonnement dans le cadre que nous avons indiqué ci-dessus pour un signal en partie connu (cas d'un système ϵ_1 qui contient une quantité d'information I_1 et qui reçoit une quantité d'information I_2 non entièrement différente de I_1), on peut raisonner de la façon suivante.

On n'a plus en fait le choix, *pour les i signaux élémentaires* déjà connus, entre N_1, N_2, \dots, N_n possibilités puisque ces signaux sont déjà exactement connus. Il ne reste donc plus qu'une seule possibilité, c'est-à-dire une certitude. Les N possibilités précédentes se ramènent donc à

$$R' = 1 \times 1 \times \dots \times N_{i+1} \times \dots \times N_p$$

et par conséquent la quantité d'information se réduit à

$$\begin{aligned}
 R' &= k \log 1 \times 1 \times \dots \times N_{i+1} \times \dots \times N_p \\
 &= k \log N_{i+1} \times \dots \times N_p \\
 &= k \log N_{i+1} + \dots + k \log N_p \\
 &= R_{i+1} + \dots + R_p
 \end{aligned}$$

C'est cette quantité R' d'information « nouvelle » dont on doit tenir compte dans l'application de la loi de combinaison. C'est également elle qui interviendra dans le calcul du rendement. On a alors

$$\eta = \frac{\text{information utile}}{\text{information totale transmise par le message}}$$

$$\eta = \frac{k \log N_{i+1} \dots N_p}{k \log N_i \dots N_p}$$

$$\eta = \frac{\varepsilon p_{i+1} \log N_k}{\varepsilon p_i \log N_k}$$

Cette formule peut s'écrire sous la forme

$$\eta = \frac{\varepsilon p_i \log N_k - i_1 \log N_k}{\varepsilon p_i \log N_k}$$

C'est-à-dire encore

$$\eta = \frac{|I_2| - |I_1|}{I_2}$$

où I_2 représente la quantité d'information totale reçue par le système et I_1 la quantité d'information contenue par ε , ces quantités étant prises en valeur absolue.

On retrouve donc exactement la formule que nous avons énoncée précédemment.

La formule d'addition de la théorie de Shannon se présente alors comme un cas particulier de la formule que nous venons d'énoncer. Dans ce cas particulier — comme nous l'avons fait remarquer précédemment — le système (de transmission) ne contient pas de dispositif de mémoire et par conséquent, ne contient jamais aucune information. On a alors $i = 0$, et l'on retrouve bien

$$R' = k \log N_i \times \dots \times N_p = R$$

NOTE SUR L'ORIGINE DE L'INFORMATION

L'origine de l'information (entité sur laquelle porte la théorie dont nous venons d'indiquer quelques éléments) ne nous paraît pas

poser de problèmes spécialement préoccupants. Tout au moins ces problèmes ne sont-ils pas plus essentiels que ceux posés sur l'origine de l'énergie et qui ne préoccupent plus sérieusement personne.

Cependant plusieurs auteurs, et en particulier des philosophes, ont voulu voir dans ce problème de l'origine de l'information un obstacle grave dressé sur le chemin de la cybernétique. Pour certains, il s'agit même d'un motif suffisant pour rejeter en bloc toute la théorie.

En fait ce problème d'origine nous paraît mal posé, d'autant plus que, dans l'esprit de ces auteurs, il se rattache à des théories qui ne se préoccupent que de la transmission de l'information (et nullement de son comportement ou de ses transformations) comme nous l'avons remarqué précédemment. Ces problèmes d'origine ont toujours été — en se plaçant au point de vue scientifique, c'est-à-dire en négligeant le côté métaphysique pour le développement d'une théorie — de faux problèmes. C'est le cas de l'origine de l'énergie en thermodynamique, de l'origine de l'expansion du monde en astronomie, de l'origine de la vie en paléontologie, etc....

Ceci est si net que toutes les sciences se développent en refusant de livrer la validité de leurs résultats à la solution des problèmes d'origine. Il doit donc logiquement en être de même pour toutes les théories portant sur l'information. C'est le contraire qui serait extrêmement surprenant.

Résoudre le problème de l'origine de l'information ne consiste donc nullement à le résoudre absolument mais simplement à établir une théorie cohérente où il n'y ait nulle part de problème d'origine non résolu, sauf à l'instant zéro. Cela revient à réaliser exactement ce qui a été fait pour la thermodynamique. Ramené à ces proportions, le problème est beaucoup moins complexe qu'on veut bien le dire parfois.

La plus grande partie des critiques qui ont été émises sur l'origine de l'information en cybernétique peuvent en effet être résumées par le syllogisme (en apparence fort correct) suivant :

- 1) La cybernétique a démontré (et ce n'est qu'une conséquence inattaquable de l'entropie) que les machines à information, de quelque type qu'elles soient, ne peuvent que diminuer le volume d'information qu'elles reçoivent ou au mieux le conserver.
- 2) La cybernétique postule que le cerveau humain n'est qu'un cas particulier de machine à information.
- 3) Par conséquent, le cerveau humain ne peut que diminuer la quantité d'information. Il est donc impossible de trouver l'origine de l'information émise par l'homme.

En fait ce raisonnement est faux, car il postule implicitement que l'homme est un créateur absolu d'information et la difficulté apparente constatée vient de là. La faille apparaît de façon évidente si l'on effectue un parallèle avec l'énergie. En transposant, on peut en effet écrire que, dans le cas de l'énergie :

1) La thermodynamique a démontré qu'une machine, quel que soit son type, ne peut que diminuer le volume d'énergie utilisable qu'elle reçoit ou au mieux le conserver.

2) Le corps humain est assimilable à une machine et soumis aux lois de la thermodynamique.

3) Par conséquent, le corps humain ne peut que diminuer la quantité d'énergie qu'il reçoit, et on ne voit plus d'où vient l'énergie qu'il fournit.

Arrivé à ce stade du raisonnement, l'erreur apparaît très nettement. Il ne viendrait en effet plus à l'idée de personne de considérer l'homme comme un créateur absolu d'énergie. Toute l'énergie qu'il fournit vient du monde extérieur (sous forme d'aliments par exemple).

Il en sera de même pour toutes les chaînes énergétiques que l'on pourra envisager et c'est justement en cela que consiste finalement l'impossibilité du mouvement (ou plus exactement du travail) perpétuel.

Si on revient alors à l'origine de l'information existant dans un « organisme » et fournie par celui-ci au monde extérieur, on voit que la réponse est exactement la même : cette information est également fournie par le monde extérieur. Dans le cas particulier de l'homme, cette information est fournie sous forme de signaux reçus par les organes de sens. Dans le cas général, ce sont les signaux reçus par les capteurs et envoyés par eux à l'organisme qui apportent cette information. La loi de conservation est respectée, tout comme pour l'énergie et il n'y a nulle part création d'information.

L'homme ou la machine ne fait alors que mettre sous une forme différente l'information qu'il a reçue. Cette mise en forme peut donner lieu à des opérations qui font intervenir l'information qui se trouve déjà à l'intérieur du système (on parlera alors de raisonnement dans le cas de l'homme, de calcul ou de « déclenchement » dans le cas de la machine).

De toute façon, l'information réémise par l'organisme considéré ne comportera jamais aucun élément absolument nouveau, quelle que soit sa représentation ou sa nouveauté apparente. Elle sera,

au maximum, égale à la somme de l'information reçue de l'extérieur et de l'information existant déjà à l'intérieur de l'organe envisagé.

Cette information intérieure peut, elle-même, avoir deux origines distinctes, en première analyse du moins. Il peut en effet s'agir de l'information inhérente à la nature même de l'organisme, information qui a été déposée par le constructeur, dans le cas d'une machine, et par les mécanismes héréditaires, dans le cas d'un être vivant. Il peut, au contraire, s'agir d'information venue de l'extérieur (par l'intermédiaire des capteurs ou des sens) et conservée grâce à une « mémoire », quelle que soit la matérialisation de cette mémoire. C'est le cas, par exemple, des réflexes conditionnés pour un être vivant ou des données de calcul pour une machine à calculer.

Dans tous les cas, l'ensemble de l'information existant dans l'organisme a, en dernière analyse, une origine extérieure et la loi de conservation n'est violée en aucun cas.

On peut aisément vérifier le bien fondé de ces affirmations en prenant des exemples simples dans la vie courante.

Considérons le cas d'un automobiliste roulant de nuit. Le système Σ que l'on doit alors considérer sera formé par l'ensemble de l'automobile et de son conducteur. Le milieu extérieur « utile » se réduira à la route et à son environnement immédiat. Dans ces conditions, la seule information reçue par les capteurs (les yeux du conducteur en l'occurrence) concerne l'état du monde extérieur au système, mais limité au faisceau des phares. Et on peut affirmer que l'action des effecteurs de la voiture (moteurs, freins, etc...) est la conséquence directe de cette information, et n'est pas autre chose que cette information mise en forme en fonction du but à atteindre et de l'information intérieure au système sur le comportement de ce système. La preuve en est que si une panne de phare survient, aucune information ne parvient plus au système qui est alors mis dans l'impossibilité d'atteindre le but fixé.

On peut comparer à ce qui se passe, au point de vue énergétique, dans une chaîne de mécanismes permettant, à partir d'une source unique d'énergie (un arbre à poulies par exemple), de mettre en mouvement nombre d'outils exécutant des travaux divers. Finalement ces travaux variés n'ont, du point de vue énergétique, pas d'autre origine que l'énergie envoyée sur l'axe.

Ce raisonnement n'implique nullement la présence d'un être humain. Prenons en effet l'exemple d'un calculateur de tir couplé à un radar de poursuite et à un canon automatique. Cet ensemble est, du point de vue informationnel, complètement autonome et se chargera d'exécuter sans intervention extérieure tous les actes

nécessaires pour atteindre le but fixé : abattre tout avion passant dans sa zone de vigilance.

Là encore, toute l'information reçue se résume à celle reçue par le radar, qui fait office d'organe des sens. Si, en effet, on « aveugle » le radar, l'ensemble, tout en restant intact, ne peut plus assurer sa fonction.

On peut d'ailleurs mettre en évidence, de façon différente, l'origine extérieure de l'information échangée entre divers « organismes ». On peut même affirmer, de façon plus précise, que l'échange d'information n'existe qu'à cause de l'ignorance dans laquelle les organismes qui reçoivent cette information, se trouvent au sujet des lois qui régissent le monde extérieur (c'est-à-dire pour chacun d'eux extérieur à eux-mêmes).

Plaçons-nous dans l'hypothèse d'un univers strictement déterministe pour simplifier les choses. Si un organisme possédait la connaissance absolue de tous les paramètres de l'univers et de leur état initial (en admettant, ce qui n'est pas évident, qu'un tel organisme ne s'étende pas alors à l'ensemble de l'univers), sa conduite pourrait être à chaque instant strictement appropriée aux conditions extérieures, sans qu'il ait besoin de recevoir aucune information venue de l'extérieur. Il serait d'ailleurs dans l'impossibilité de recevoir une information utile, car, si l'on applique la formule du rendement informationnel que nous avons énoncée ci-dessus, on voit que ce rendement ne pourrait être que nul à tout instant, toute information reçue par l'organisme y étant déjà contenue.

En fait, une telle connaissance de toutes les lois est impossible pratiquement, sinon théoriquement. C'est pourquoi tous les organismes existants ont adopté soit une approximation de cette méthode, soit une autre solution.

Dans cette seconde solution, la quantité d'information, contenue de façon explicite ou non dans l'organisme, a été réduite (vraisemblablement au minimum compatible avec les performances de cet organisme) et ce sont les capteurs, couplés, tant au point de vue énergétique qu'informationnel, avec le monde extérieur, qui informent à l'organisme ce qui se passe dans son monde extérieur. (Il faut entendre par là le monde extérieur qui intéresse l'organisme et dont la connaissance lui est accessible grâce à ses capteurs). On obtient ainsi, à chaque instant, une vue instantanée du « monde extérieur utile » et il est possible d'en déduire, par l'emploi de quelques règles simples (« héréditaires » ou apprises), le comportement le plus adapté à l'instant immédiatement suivant. Il y a donc

correction continuelle du comportement grâce à un feed-back du type le plus classique.

En somme, alors qu'un organisme idéal connaissant tous les paramètres et toutes les lois de l'univers aurait à chaque instant une conduite strictement adaptée, un organisme recevant de l'information extérieure ne sera jamais qu'approximativement adapté : la courbe résumant son comportement ne pourra qu'approcher la courbe idéale décrite par le premier organisme. La précision de l'approximation est alors fonction des caractéristiques des capteurs et des lois à appliquer.

Il est du plus haut intérêt de remarquer que les deux solutions ont été essayées aussi bien par l'évolution pour les êtres vivants que par l'homme pour la réalisation de ses machines. Dans les deux cas, c'est la seconde solution qui a triomphé malgré une précision moindre au départ. Elle présente en effet une souplesse beaucoup plus grande qui lui permet, à complexité égale, de résoudre un nombre de problèmes nettement plus important, et par conséquent, de mieux s'insérer dans le milieu extérieur.

Dans le cas des organismes vivants, c'est l'instinct qui correspond à la première solution, mais il ne peut s'appliquer que sur une portion limitée du monde extérieur. Les actes effectués atteignent un degré de complexité et de précision élevé, mais si un obstacle imprévu s'interpose il n'existe aucune adaptation possible. Il n'y a pas non plus de possibilités d'apprentissage et de perfectionnement.

L'intelligence correspond, au contraire, à la seconde solution et, si son comportement est moins satisfaisant au début, elle laisse la possibilité d'une adaptation et d'un développement grâce à l'apprentissage.

Dans le cas des machines, la première solution correspond aux machines classiques sans feed-back, la deuxième aux machines asservies, avec feed-back. La première solution a abouti à des machines très complexes et très précises, mais qui, finalement, ont été condamnées par la quasi-impossibilité où l'on se trouvait de les adapter à un changement, même minime, du résultat à obtenir. La deuxième, base de toute l'automatisme moderne, connaît actuellement un triomphe qui semble devoir être total.

On voit donc qu'en dernière analyse, toute l'information existante provient de la totalité du monde extérieur et qu'il n'y a, à aucun moment, création d'information.

Cette démonstration a une conséquence évidente : dans ces conditions, la quantité globale d'information utile de l'univers ne peut que diminuer au cours du temps. Cette conclusion est parfaite-

ment d'accord avec l'un des résultats les mieux établis de la physique : l'augmentation continue de l'entropie. En effet l'information se présente comme l'inverse de l'entropie et les deux phénomènes sont parfaitement parallèles.

Cette relation montre également que l'information utilisable contenue dans l'univers a la même origine que l'énergie utilisable et corrobore notre affirmation de l'identité profonde de l'énergie et de l'information.

The value of the application of cybernetic principles to physiological regulatory mechanisms

by H. LABORIT and B. WEBER, ⁽¹⁾

*Physiobiological Research Section
of the French National Navy (Paris)*

The application to physiological phenomena of the concepts and investigational means offered by cybernetics should greatly increase the understanding of these phenomena. We would like to demonstrate first that the use of a clear method of representation will furnish the biologist with a useful working tool. It will force him to set down in a precise manner the problems he intends to solve and will enable him to integrate and visualize many autoregulated functionings, the complexity of which often leads to confusion. The cybernetic representation will be thus called on to become in physiology what algebra is in mathematics.

GENERAL NOTIONS AND CONVENTIONAL REPRESENTATIONS

An "effector" is shown represented by a circle (fig. 1); its "effect" (or one selected from all the effects produced by this effector) is represented by an arrow originating from the circle (E); its "factors" are represented by arrows leading to it (A, B).

Such a system may contain a mechanism that is sensitive to the variations of the effect and that is capable of acting on one or more factors in accordance with these variations. In such a case, the factor(s) become a function of the effect, which will also remain a function of the factors that control it.

Therefore, an autoregulation exists since the variations of the effect caused by the variations of the factors will, in turn, influence the latter. The existence of such a "feedback" mechanism acting on one of these factors is represented in figure 1 by a curved arrow

(1) This article was sent to the secretariat of the review on the 23 th of July 1959.

originating from the effect arrow and leading to the arrow of the factor(s).

Such a mechanism permits the effect to liberate itself partly from the influence of the factor variations that determine it. Thus the feedback modifies to a certain extent the variations of these factors.

The purpose of an effector is to achieve an effect of a particular intensity. If this purpose is reached and the intensity maintained, it can be said that the effector is functioning in "constancy". As we shall see, this is most frequently the case in physiological effectors (fig. 2, B).

If the purpose of the effector is to achieve an effect of a maximum intensity, which is never reached and can even tend toward infinity, it is then said that the effector functions in "tendency" (fig. 2, A).

An effect can be "positive" or "negative". Similarly, the effector can be either positive or negative, insofar as its variations influence the variations of the effect in the same or in the opposite direction.

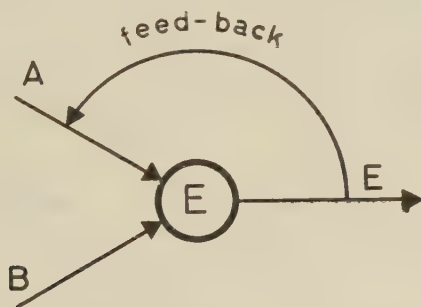


Fig. 1.

Representation of effector, its effect and its factors.

A feedback can also be either positive or negative. An effector functioning in constancy is properly regulated if it is subjected to a negative feedback. In the opposite case it will lose its regulation.

Conversely, the effector functioning in tendency is under the influence of a positive feedback. In the organism, the tendency is generally limited to a certain value; the functioning of the kidney tubule cells is a good illustration of this.

In figure 1, it can be seen that the feedback, which originates from effect E, will influence the factor with a certain delay. Hysteresis results therefrom. Similarly, the modification introduced in factor A will be felt by E only with a certain delay, which is called the efficiency delay. These two concepts explain, for example, why an

effector in constancy maintains its equilibrium around an ideal value, which is never stably reached, and why the effect goes through oscillating values. Here again we find the harmonious postaggressive oscillating reaction. We would now like to explain why we describe a disharmonious postaggressive oscillating reaction in physiopathology. The reason is that an effect can be produced only if the intensity of its factors remains within given limits.

An effect depends on many factors of variable intensity. It is a causative function. The limits of a causative function are formed by the ensemble of the values that can be assumed by all these variables without resulting in the interruption of the effect.

In physiopathology it can thus be assumed that the effect will not be produced if the eventual variations of the environment lead to variations of the intensity of certain factors forcing them outside of their limits. We shall see that if the variations of the environment force the factors responsible for the regulation of the equilibrium of a complex organism outside of their limits, the organism will alter its *modus operandi*.

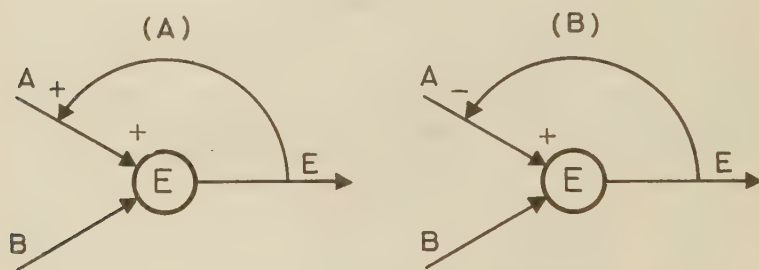


Fig. 2.

(A) Positive feedback, functioning in tendency.

(B) Negative feedback, functioning in constancy.

APPLICATIONS TO PHYSIOLOGY

With the help of these extremely simplified notions, we must first try to discover the nature of the finality of a living organism. Since the work of Claude Bernard, it has been repeated that the finality is the maintenance in the internal environment of the constancy of the conditions necessary for life, in other words, of Cannon's homeostasis. We must face the fact that this constancy disappears in physiopathology. It would be too time-consuming to explain here what leads us to accept that the maintenance of this constancy

(which obviously calls into play autoregulated systems) is only the means to achieve the true finality. This finality is the maintenance of the degree of organization of the living matter in spite of and against the tendency toward leveling, toward entropy. To be accurate, the study of the cybernetic mechanisms that achieve it should therefore begin at the level of the protein molecule, then progress toward that of the unicellular organism. As far as we know, such a study has not yet been attempted. As far as we are concerned, we shall begin to apply these concepts at the cell level, and we shall accept that the first "finality" in a complex organism is the maintenance of the cell organization.

CELL AUTOREGULATION

We have previously described the mechanism of cell autoregulation.

When it functions under aerobic conditions, the cell as an effector can be schematically represented (fig. 3) as having for effects the liberation of H ions and of CO_2 . The factors are represented by dehydrogenation and decarboxylation processes. The selector mechanism is represented by the membrane, which controls permeability.

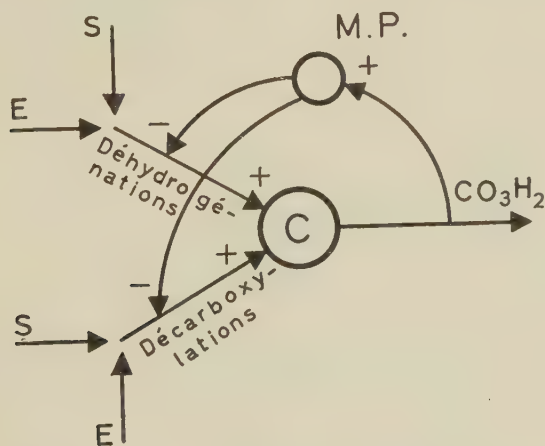


Fig. 3.

Cell autoregulation functioning under aerobic conditions.

C: cell.

E: enzymes.

M. P. : membrane potential.

S: substrates.

The more the metabolism increases (dehydrogenations plus decarboxylations), the greater is the liberation of H ions and CO_2 and the more the membrane potential increases. This increase results in a proportional decrease of membrane permeability, leading then to a proportional decrease in membrane potential, and so on. This is a typical example of a constancy functioning.

PHYSIOLOGICAL ORGANIC AUTOREGULATION

The extracellular environment or the ratio of the concentrations in BHCO_3 and H_2CO_3 of this environment $\left(\frac{[\text{BHCO}_3]}{[\text{HCO}_3\text{H}]}\right)$ may be taken

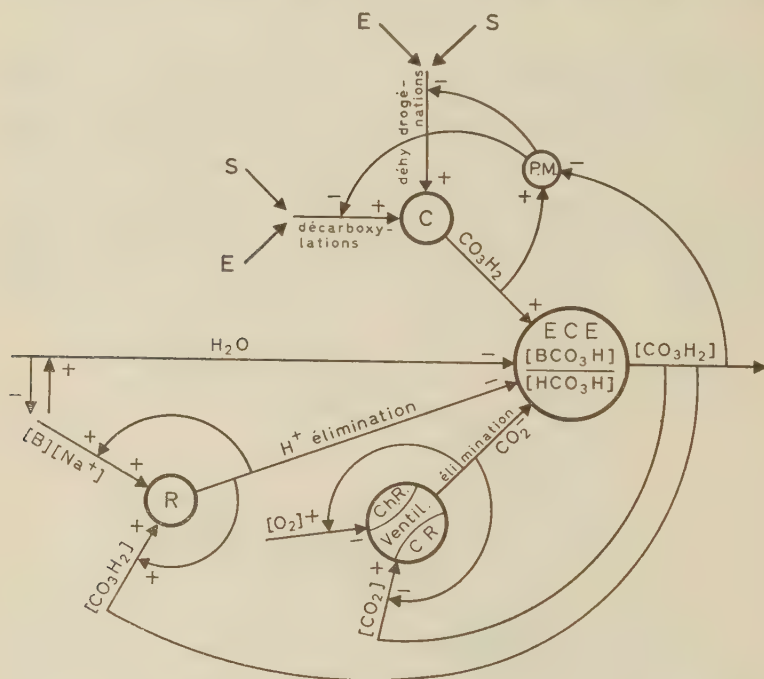


Fig. 4. - Effect on the extracellular environment effector of its H_2CO_3 concentration.

- C : cell.
- E : enzymes.
- P. M. : membrane potential.
- S : substrates.
- ECE : extracellular environment.
- R : kidney tubule cell.
- Ventil. : ventilation.
- Ch.R. : chemoreceptors.
- C. R. : respiratory centers.

arbitrarily as the effector, and the maintenance of the concentration in H ions may be taken as the purpose of this effector.

Under aerobic conditions, the factor that will cause a positive variation of this concentration is the liberation by the cell of H ions and CO_2 . Figure 3 will therefore terminate at the next effector, "the extracellular environment", and the effect of this cell will become the positive factor on the extracellular environment. The effect selected for this latter will be its H_2CO_3 concentration (fig. 4). When it increases, the cell membrane becomes depolarized. We shall therefore, draw a negative feedback starting from this effect and leading to the membrane potential. We have seen, however, that membrane polarization acts negatively on metabolic intensity. Two negative actions result in a positive one. The increase of the H ion concentration — acidosis — as long as it remains within physiological limits, will increase metabolism; a positive feedback is involved, and therefore the system will function in tendency. If one or more negative factors did not permit equilibration of the H_2CO_3 concentration of extracellular fluids under constancy conditions, the cell organization would be very rapidly destroyed since we have just seen that the maintenance of this organization is possible only under constancy conditions. This results from the fact that in figure 3 we have selected as the effect on the cell effector the amount of H_2CO_3 liberated and that in figure 4 we have taken as the effect on the extracellular environment effector its H_2CO_3 concentration. We have progressed from the cell level to that of the organism.

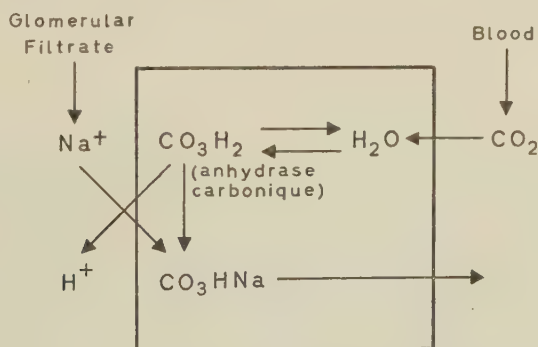


Fig. 5. — H ion excretion mechanism of the kidney.

The negative factors are the kidneys and the lungs, which assure the evacuation of CO_2 and H ions outside of the organism. The functioning of the kidney can be schematically represented. For simplicity's sake we shall limit ourselves to only one of its H ion excretion mechanisms (fig. 5).

It is known that the cell tubule can produce H_2CO_3 from CO_2 and the H_2O of the plasma through the action of the carbonic anhydrase. The H_2CO_3 traps the sodium filtrated at the level of the glomeruli and forms NaHCO_3 , which then returns into the circulation. The H ion is then liberated into the urine. This mechanism can be represented as in figure 6.

A positive feedback — a functioning in tendency — is therefore involved, which will lead to reintegration into the organism of that much more CO_2 and Na so that there will be more H ion excreted in the urine. Conservely, the more the plasma concentration of CO_2 and Na increases, the more H ions will be excreted (combined with strong anions), but the greater also will be the CO_2 and Na retention. This mechanism can be limited only by parallel lowering of the H ion concentration in the organism. We shall see later in a few examples that this hypothetical representation is supported by facts that are thereby clarified.

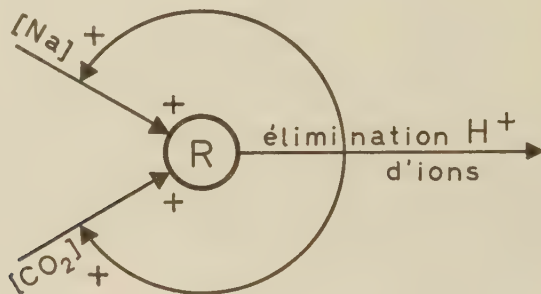


Fig. 6.

Mechanism of liberation of the H ion the urine.

It should be noticed that the more sodium is retained, the more water is also retained. An “interaction” between two factors acting on the “extracellular environment” effector therefore exists since the more water is retained the more the H_2CO_3 concentration of the extracellular fluid is decreased. This leads to a greater decrease of the CO_2 and Na concentrations. A third negative factor should, therefore, be added to the regulations of the H_2CO_3 concentration of the extracellular fluid: the extracellular water. Its regulation will be considered later.

In the lungs (fig. 4), the ventilation can be taken as the effector, and it is controlled by the respiratory center (CR) — particularly sensitive to the blood CO_2 concentration $[\text{CO}_2]$ — and the chemore-

ceptors (ChR) — particularly sensitive to the blood O_2 concentration $[O_2]$ (fig. 4). The higher the O_2 concentration, the more the chemoreceptors are inhibited. This is a negative factor on the ventilation. On the contrary, the more the CO_2 concentration increases, the more the respiratory centers will be stimulated (it is understood that this holds true under physiological conditions, that is to say, when the variable factors remain within certain given limits), and this is a positive factor on the ventilation. When ventilation increases, more CO_2 is excreted and less is found in the blood : a negative feedback on a positive factor ; thus there exists a functioning in constancy, which maintains the equilibrium of the effect. Similarly, when the ventilation increases, the oxygenation improves ; now a positive feedback on a negative factor on the ventilation is involved ; therefore, here also the functioning is in constancy.

Let us observe in passing some facts made obvious by this cybernetic representation. Let us suppose, for example, that the excitability of the respiratory centers is depressed by an anesthetic, and that our patient is made to breathe a gas mixture rich in oxygen. It is obvious that if the positive factor on ventilation is absent or inadequate, the increase of the value of the negative factors will lead only to a deeper depression of respiration, whereas when the respiratory centers are intact, the respiratory depression that could result from the increase of oxygen concentration of the inspired air is immediately compensated by the increase of the pCO_2 . An opposite line of reasoning can be followed when the decrease in the partial pressure of the oxygen in the inspired air is concerned. If it causes a hyperventilation, it will lower the CO_2 and depress the activity of the respiratory centers. Since these centers are comparatively more sensitive, anoxic anoxia causes but little ventilatory compensation.

It can, therefore, be seen that the equilibrium in the concentration of CO_2 of the extracellular fluid can be expressed as the maintenance of the equilibrium between the cell production of H_2CO_3 on the one hand — a positive factor — and as the excretion of CO_2 by the lungs and of H ions by the kidneys on the other hand — both negative factors. The water concentration, which is related to that of sodium and therefore of CO_2 , is also a negative factor acting on the H_2CO_3 concentration.

It should finally be added that the positive feedback that regulates the effect, i. e., the liberation of H ions and CO_2 by the tissues and their excretion out of the organism by the emunctories, originates from this same effect. Indeed, this is because the organs responsible for the excretion of the H ions and CO_2 are also formed by cells, the

metabolism of which is autoregulated and involves the same mechanism found in other tissues. The increase of H_2CO_3 concentration in the extracellular environment results in an increase of ventilation. Such is the case, for instance, in muscular exertion because it depolarizes membranes and renders them more permeable, increases cell metabolism through a positive feedback mechanism, and increases simultaneously the work of the cardiovascular and renal systems. The neurones of the respiratory centers become depolarized and therefore stimulated, and ventilation is increased. A positive feedback on a negative factor is involved at the level of most tissues and leads to a functioning in tendency that perpetuates the disturbances. But a positive feedback on negative factors is involved at the level of the organs mobilizing and excreting the H ions and CO_2 and leads therefore to a functioning in constancy and to a correction of the disturbances.

Let us consider now as the effect of the extracellular environment its sodium concentration $[\text{Na}^+]$. It can be expressed as osmotic pressure (fig. 7) if we agree to overlook temporarily the other elements that enter into play in establishing its value. The sodium concentration has a positive factor, which is the amount of sodium contained in the extracellular fluid, and a negative factor, which is the amount of water present in this same environment. It is admitted that the over-all value of these two factors depends on the secretion of aldosterone by the adrenals. Indeed aldosterone is responsible for water and sodium retention. This retention, however, inhibits the aldosterone secretion. This is a negative feedback and a functioning in constancy that probably bring into play the still much debated volume receptors.

Osmotic pressure (Na^+), however, acts negatively on the osmoreceptors, which are stimulated by dilution. These osmoreceptors act on the posterior lobe of the pituitary to inhibit it, and again a negative action is involved. The pituitary secretes an antidiuretic hormone, which also acts negatively on diuresis, and in turn, finally acts negatively on the amount of water in the extracellular environment. A positive feedback on a negative factor is, therefore, involved leading to the functioning in constancy of the osmotic pressure (Na^+).

The illustrations can now be combined by using the osmotic pressure in figure 7 (since we have stated that we identified it with Na^+) as the factor of the H^+ elimination at the level of the kidney tubule cell, and by taking as the origin of the H_2O negative factor of figure 4 the H_2O factor of figure 7.

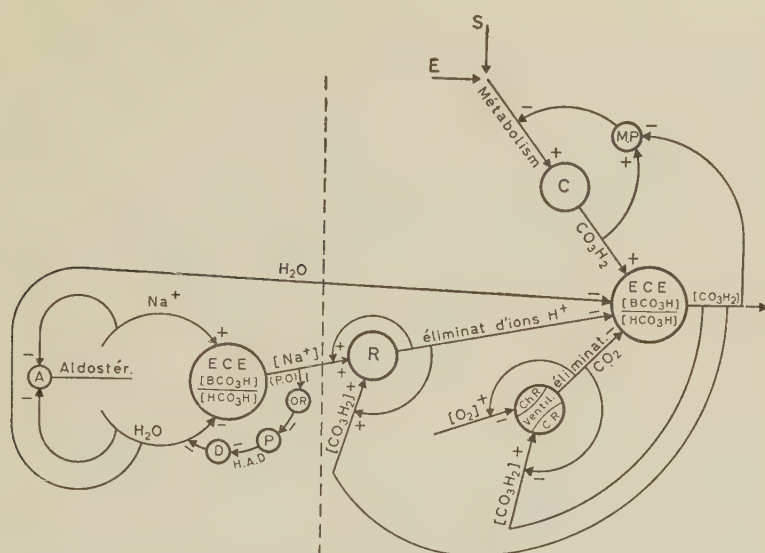


Fig. 7.
Osmotic pressure.

- C : cell.
M. P. : membrane potential.
E.C.E. : extracellular environment.
R : kidney.
Ventil. : ventilation.
Ch.R. : chemoreceptors.
C.R. : respiratory centers.
A : adrenals.
O. R. : osmoreceptors.
P : pituitary.
D : diuresis.
H.A.D. : antidiuretic hormone.
P.O. : osmotic pressure.
Aldostér. : aldosterone.

HOW TO APPLY THE PRECEDING CONCEPTS

First, it appears that all living phenomena under physiological conditions can be summarized as the conversion within the cell of H molecules bound to the substrates supplied by diet into H ions and electrons accepted by oxygen. These are then carried to the emunctories and excreted. Another observation that must be made is that the transfer of hydrogen through the organism leads to a relative of "bases", in particular of sodium and of water, since the more the kidney excretes H ions, the more it retains sodium and

water. This phenomenon, however, will depress aldosterone secretion. Consequently, the quantitative water and salt homeostasis will be restored.

Moreover, this representation seems to bring to light certain correlations between pH and osmotic pressure, which at first sight would not be quite evident. Indeed, the tendency to acidosis will be accompanied by sodium retention and the osmotic pressure will rise. The positive feedback that acts through the osmoreceptors will lead to water retention. Since the total amount of water and sodium will rise, the secretion of aldosterone will be depressed and the osmotic pressure controlled. Conversely, metabolic alkalosis will be regulated, on the one hand, by ventilation that will be depressed and, on the other, by the decrease of the H ion and the increase of bicarbonates excreted in the urine. Since the osmotic pressure decreases, the stimulation of the osmoreceptor will increase diuresis and osmotic pressure will again be controlled.

The most interesting fact, however, concerns the mechanism of the disturbances that are likely to appear if an organism is maintained under constant ventilation. By suppressing one of the main regulating systems of the H_2CO_3 concentration of the extracellular fluid, the equilibrium can be maintained only if the metabolic intensity remains adapted to this stable ventilation. Starting from this type of equilibrium, which can be set initially, the diagram will show that if cell metabolism is disturbed by any type of pharmacological means, the organism cannot function any more in "constancy" but only in "tendency" and that a metabolic disturbance will amplify itself very rapidly through a positive feedback. This concept seems to us to be of major importance. Indeed, we feel that it is not an exaggeration to state that all of pharmacology must be studied again under constant ventilation since, on an animal under equilibrium, the introduction into the organism of a pharmacological agent will be accompanied by a rapidly dampened oscillating reaction if the disturbance that the agent causes on cell metabolism does not go beyond the limits within which autoregulation remains possible. The main metabolic cell disturbance will thus be hidden, whereas constant ventilation will cause its progressive amplification even when small doses are used. According to this same principle, ligature of the ureters can also furnish useful information.

This schematic representation is obviously extremely simplified in order not to overcomplicate from the start the description of the application to physiology of the cybernetic representation. Innumerable effects, effectors, factors, and feedbacks, the importance of

which cannot be underestimated, have not been taken into consideration. In this schematic representation, we have not attempted to integrate endocrine regulation mechanisms or, for instance, the cardiovascular system, the role of which is considerable in the regulation of the equilibrium of the H ion of the extracellular environment. On the basis of the preceding findings, nothing prevents their integration. We have studied this with profitable results, but the presentation of such an integration cannot be included into an introductory article. It seems more useful to demonstrate how these principles can be put to best advantage.

APPLICATION TO PHYSIOPATHOLOGY

It was previously mentioned that an effect can be produced only if the intensity of the factors remains within certain limits. This is true in physiology where hysteresis is responsible for the oscillating nature of the reaction that remains harmonious. The intensity of these factors depends on the environment, and it is obvious that the maintenance of the physiological state is possible only within certain narrow limits of variations of this environment. At the metabolic level, the physiological state is accompanied by aerobic respiratory processes. When the variations of the environment go beyond the limits within which the factors can give rise to effects that control the variations of these factors through a feedback mechanism, experience proves that the organism can still recover its equilibrium. But experience has taught us also that, in such cases, the autoregulated mechanisms that are called into play are not those that assure the maintenance of this equilibrium under physiological conditions. Physiological feedbacks are indeed unable to maintain the factors within their limits. Identical observations are very well known in cybernetic machines. In such a case the machine must switch from one type of *modus operandi* to another. The living machine does the same thing but one point must be made clear. A change in the *modus operandi* involves the means since the finality remains unchanged. This is the same finality mentioned previously : the maintenance of the degree of organization of the living matter. Under physiological conditions the means used to maintain this degree of organization is the maintenance of the constancy of the internal environment. Under physiopathological conditions, however, biological observations have long demonstrated that this constancy is not immediately restored : the reaction becomes disharmonious. The new temporary *modus operandi* uses flight or struggle to achieve

finality. Stable splanchnic and cutaneous vasoconstrictions, acidosis, hemoconcentration, hyperglycemia, hepatic and renal anoxias leading to hyperammonemia and anuria, etc., confirm the temporary departure of the organism from the over-all homeostasis to favor the maintenance of this homeostasis at the level of certain privileged areas, which make possible flight or struggle: nervous system, heart, lungs and muscles. If this change in *modus operandi* permits the individual to flee or suppress the aggression imposed by the environment in a short enough time so that the disfavored organs do not suffer irreversibly, the finality of the organism — that is to say, the maintenance of its degree of organization — is thereby achieved. The return to homeostatic equilibrium will take place more or less rapidly, but if not, immediate or delayed death will result.

What is the purpose that should be selected to define the new *modus operandi* followed by the organism to assure its finality under these abnormal environmental conditions? Observations show that the maintenance of the constancy of the H ion concentration in the extracellular fluid is not the purpose sought to assure this finality. On the contrary, the hypoxia consecutive to the local circulatory changes leads to an anaerobic form of life more or less marked at the level of certain organs; this leads to an increase of glycolytic processes at the expense of respiratory processes and to the liberation by these organs of strong acids more dissociated than H_2CO_3 . But these “strong” acids should be excreted by the kidneys; however, their vascularization is interfered with and diuresis stopped.

The only purpose that seems logical is the intensity of the energy liberated by the organs that assure freedom of motion with respect to the environment. It is indeed this energy liberation that, through the flight or struggle it permits, preserves the finality — the maintenance of the organization of living matter — by avoiding or suppressing the aggression, i. e., the contingent cause that forced the factors outside of their limits. In such a case, it is conceivable that organs, such as the kidneys, that control negative factors of the H ion concentration of the extracellular environment have their functioning inhibited. It can also be understood that the organism will not function in constancy but in tendency so long as the disappearance of the contingent cause responsible for the switch in *modus operandi* is not eliminated. The energy liberated by the organs assuring autokinesis will increase progressively as long as the abnormal factor keeps on acting and total depolarization is not achieved.

The role that training can play in many borderline cases can finally be understood. The result of training will be indeed to broaden the limits of the variables, to widen their boundaries, without *ipso facto* causing a cessation of the effect and necessitating a change in the *modus operandi*.

Ammonia metabolism can be taken as a typical cybernetic representation of physiopathological phenomena (fig. 8).

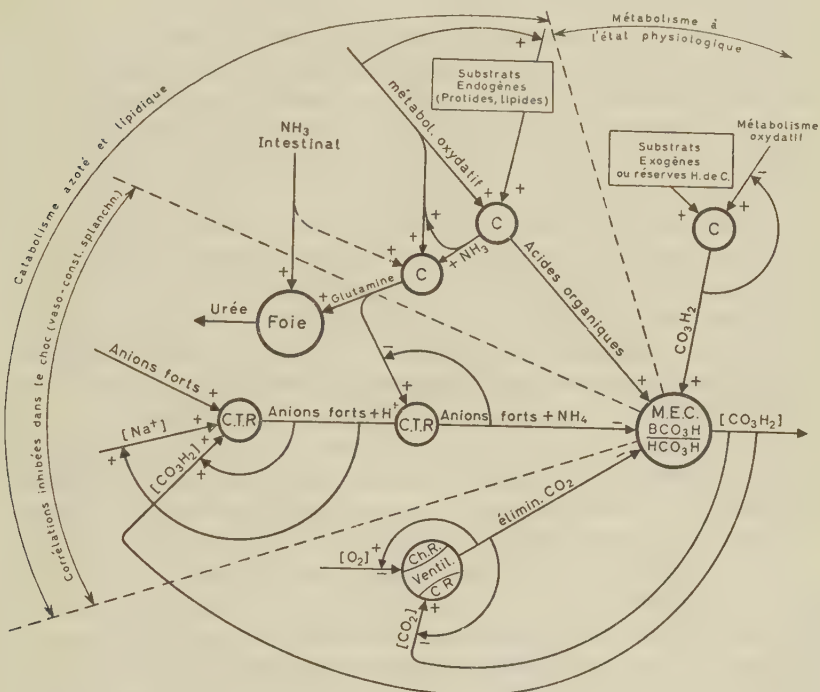


Fig. 8.

Ammonia metabolism.

Anions forts : strong anions ;

Catabolisme azoté et lipidique : nitrogen and lipid catabolism ;

Corrélations inhibées dans le choc (vaso-const. splanchn.) : correlations inhibited in shock (splanchnic vasoconstrictions) ;

Métabolisme à l'état physiologique : metabolism under physiological conditions ;

Substrats endogènes (protides, lipides) : endogenous substrates (proteins, lipids) ;

Substrats exogènes ou réserves H. de C. : exogenous substrates or carbohydrate reserves.

In the organism under physiological conditions, ammonia metabolism appears as of secondary importance since the healthy liver can metabolize intestinal ammonia and the exogenous supply of carbohydrates can assume an essential part of cell metabolic expenses. Thus the H_2CO_3 concentration of the extracellular environment is easily maintained.

As soon as the organism must have recourse to lipids, proteins, and neoglucogenesis to ensure its energy supply, the metabolism leads to the formation of organic acids and ammonia, as can be seen in fasting. The liver and muscle glycogen reserves are practically exhausted after 12 hours of fasting. This is also observed after vomiting, anesthesia, febrile conditions, and thyrotoxicosis. Oxidative metabolism is frequently disturbed here and leads to a worsening of the acidosis. In diabetes, finally, the inability to utilize carbohydrates is at the origin of the acidosis.

In summary, these cases are characterized by nitrogen and lipid catabolism, caused by the lack of carbohydrates or the inability to utilize them in the oxidative cycle.

The overproduction of NH_3 by the cell will induce an overproduction of glutamic acid to convert NH_3 into glutamine and thus detoxify it. The increase demand in glutamic acid is associated with an increase demand in α ketoglutaric acid and consequently with a speeding up of all the steps in the tricarboxylic acid cycle. The oxidative metabolism is thus suddenly increased. But it can be increased only at the expense of endogenous substrates resulting in an increase of protein catabolism and of NH_3 formation. The cell works thus in tendency and no longer in constancy. The exhaustion of the tricarboxylic acid cycle may take place, and at the level of the central neurons it would then be manifested by coma, according to Bessman. It is only when the tricarboxylic acid cycle proves to be unable to perform the NH_3 transamination into glutamine that ammoniemia increases. Hyperammoniemia appears, therefore, above all as a sign of deep disturbances of the oxidative metabolism.

Glutamine will be carried to the liver and the kidneys, assuming in this case that they function normally. In the liver, the ornithine cycle will lead to the formation of citrulline and urea. In the kidney, strong anions will be excreted to compensate for the liberation by the cell of organic acids, which have replaced HCO_3 in the ratio $\text{BHCO}_3/\text{HHCO}_3$.

The CO_2 will be excreted by the lungs and the Na reabsorbed by the tubule cells in exchange of an H ion. As is well known, the latter

will combine with a strong anion to yield a strong acid, which, in turn, will form an ammonia salt at the level of the tubule cells by taking away NH_3 from glutamine.

The blood returning to the liver after passing through the kidney will contain that much less glutamine so that there will be more strong acids to excrete to compensate for the acidosis resulting from the liberation of organic acids by the cells. The liver will have that much less material available to synthesize urea, all other factors remaining unchanged as far as the intestinal NH_3 is concerned. The coefficient of ureagenetic deficiency thus evolves around an essential function, that of the maintenance of the acid-base balance. In such cases, hypochloremic types of acidosis are observed.

In all cases it can be well understood that it is important to give a massive carbohydrate supply capable of limiting maximally the nitrogen catabolism, which generates NH_3 and organic acids. Insulin will facilitate carbohydrate intermediary metabolism.

If the hepatic function is seriously disturbed (as in cirrhosis, hepatitis, vena cavaportal anastomosis), the excretion of NH_3 will rest only on kidney function and the excretion of strong acids. In such a case again, it is urgent to prevent the nitrogen catabolism and to limit the formation of intestinal NH_3 . The considerable danger of certain therapeutic agents, such as carbonic anhydrase inhibitors (acetazolamide), can be understood. They favor bicarbonate loss by the kidney, hinder the excretion of strong acids, and make the acidosis and thus nitrogen catabolism even worse. But above all, because they limit NH_3 renal excretion, they force the cell to speed up metabolism, which rapidly ends in exhaustion and coma.

The liver and the kidneys can both be damaged. Moreover, the vasomotor reaction to various aggressions is always accompanied by splanchnic vasoconstriction and by the functional inhibition of these organs: the situation is no longer chronic but becomes acutely pathological as observed in the various shock states as well as in burns. The acid-base balance can no longer be maintained except by ventilation, therefore very unsatisfactorily and temporarily. Cell metabolism then will function in tendency since it is practically under the influence of only a positive feedback. The change in the *modus operandi* (flight to struggle) must rapidly comply with the organic finality, otherwise death by exhaustion will be rapid. Here again, the importance of a massive glucose supply as well as of the suppression of the splanchnic vasoconstriction by neuroplegia is obvious. It also becomes clear that it is important to sterilize the intestine with antibiotics since the flora of the intestine generates

NH₃, which the anoxic liver cannot detoxify ⁽¹⁾. The entire classical postoperative nitrogen catabolism is condensed here, since the adrenergic system and the pituitary-adrenal axis play an essential correlation role between splanchnic vasoconstriction and cell nitrogen catabolism.

In all the cases cited, provided that intensity and the duration of the aggression as well as the seriousness of the lesions are controlled, it is possible to understand the importance of an imposed restoration of homeostasis and, in the first place, of circulatory homeostasis (suppression of the vasoconstriction, which has become useless and is a factor of metabolic acidosis). The flight or the struggle has lost its meaning and the switch in *modus operandi* is no longer necessary. Similarly, it is possible to understand the importance of maintaining the balance between formation and excretion of H ions and CO₂ and in some cases, therefore, of decreasing the metabolic intensity by cold when the mobilization and excretion of H ions and CO₂ are threatened. Neuroplegia and artificial hibernation, accordingly, find some additional supporting arguments. If the restoration of homeostasis is imposed, however, the physiological *modus operandi* must be preserved, and vice versa. Thus, in cases of hemorrhage, a neuroplegia without restoration of the blood mass is as illogical as a transfusion without neuroplegia.

This is all merely a more or less hypothetical and simple interpretation of facts. Cyberneticians have taught us the importance of "models" in the understanding of a complex mechanism. The change in *modus operandi* that we have just described in the complex organism, passing from physiology to pathology, is illustrated by Ashby's homeostasis. When a particular determinism does not allow the obtaining of the effect of its finality, the latter will reorganize its *modus operandi* and seek new equilibrium conditions.

This very schematic synopsis is intended only to draw the attention of the physician and the biologist to the possibilities offered by cybernetic representation and method of reasoning, and we have demonstrated that this method can help us :

(1) set down correctly and diagrammatically the data of a problem in order to arrive at its solution clearly ;

(2) integrate and thus correlate factors the complexity of which often renders a synthesis difficult, and follow in a simple fashion their more or less far-reaching consequences ;

⁽¹⁾ We observed that the hyperammonemia that is associated with the shock episode in severe burns is brought back to normal by plain neuroplegia.

(3) bring to light certain facts and the importance of certain experimental protocols ;

(4) define and satisfy better the internal finality of complex organisms such as the human body.

We hope that this unavoidably fragmentary first attempt will incite those interested in this field to develop the method. It seems evident that it could help considerably in the understanding of the extremely complex autoregulated mechanisms that assure the maintenance of the organization of living entities in the midst of variations of environmental contingencies.

BIBLIOGRAPHY

- ASHBY, W. R., *Introduction à la cybernétique*. Dunod, Paris, 1958.
BESSMAN, S. P. AND BESSMAN, A. N., *J. Clin. Invest.*, **34**, p. 622, 1955.
BESSMAN, S. P. AND BRADLY, J. E., *New England J. Med.*, **263**, p. 1143, 1955.
COSSA, P., *La cybernétique*. Masson et Cie., Paris, 1955.
LABORIT, H., *Bases physiobiologiques et principes généraux de réanimation*. Masson et Cie., Paris, 1958.
LATIL, P. (de), *La pensée artificielle*. Gallimard, Paris, 1953.
VENDRYES, P., *Vie et probabilité*. Albin Michel, Paris, 1942.
-

Observations on the role of business research as an aid to managers

by M. L. HURNI,

*Manager, Operations Research & Synthesis Consulting Service,
General Electric Company (New York)*

In speaking of the role of business research as an aid to managers, I am going to make an assumption. This assumption is simply that for the foreseeable future, most industrial businesses will be constituted pretty much as they are now. They will be loosely coupled systems in which many people will continue to play a major role in sensing, deciding, regulating, adjusting, innovating.

It is further assumed that such sensing and deciding will be done in many areas of a business, that it will be done and can only be done on the basis of what people so situated know and believe, and that it can only be done voluntarily.

In short, I intend to talk about applied research — that is applied business research *now*, not in an environment like that described in George Orwell's *1984* ⁽¹⁾ or in Vonnegut's *Utopia 14* ⁽²⁾ where the world round about has somehow become static and the task is principally maintenance of the status quo.

In this year of 1960, it is rather difficult to imagine the president of a corporation being content to sit in his "control room" watching indicating instruments or reading "print-outs" from a computer that has been programmed to describe a business for from now to eternity by his staff of scientists, and making decisions accordingly, if decisions still need to be made, that directly

(1) ORWELL George, *1984*. New American Library of World Literature, Inc., 1954.

(2) VONNEGUT, Kurt Jr., *Utopia 14*. Bantam Books, New York, 1952.

control sales activity, manufacturing systems, direct labor, to mention a few.

In the first place, it is doubtful if such a man would and should be of such sedentary disposition.

Secondly, it is doubtful whether he knows or can state the objectives of his business so precisely for the long haul as to make this possible.

This is not a criticism of a man so situated, but rather is based upon his deep intuitive understanding that a business enterprise is not a static thing, but something that is always "becoming". It is subject to change and is also intended to create change. It is an organism rather than a machine.

It is also based upon an intuitive understanding that objectives are not simple; that the implications of words like maximize or optimize are not sufficiently descriptive of the situation — convenient and useful as they may be in the generation and solution of idealized problems.

One might say that in the 1960's this is not possible anyway. It has nothing to do with the nature of the subject matter of research; namely, the business-enterprise, but rather with the state of development of research.

It is just a case that the tools available now and in the foreseeable future are not sufficiently powerful to treat with complete rigor the complexity inherent in a total business. It is simply that more work, and more detailed work needs to be done on smaller pieces.

So we go to the other extreme, the foreman or supervisor and the smaller component of which he is a subfunctional manager. In 1960, it is also difficult to visualize this individual abdicating to a set of decision rules provided by a research team while he is still accountable for results from the component — on the assumption that such decision rules can be provided by applying linear programming or game theory, for example.

It is also difficult to visualize such an individual as willing to think only of considering his own situation without regard for his obligation to others involved in the flow of materials, processed parts, information, for example — or for that matter, of the obligations of others to *him* if the rules are to work as expected.

Although this description of the situation in 1960 might be greatly oversimplified, it does depict a dilemma. That dilemma is concerned with making business research effective and of high

impact within a business enterprise. I would not be content to draw conclusions from this simplified statement were it not for the fact that this dilemma also seems to exist in real life.

Recall I am speaking of applied research in industrial business, not pure research for the sake of acquiring knowledge, not applications in the military area, nor in airborne or space technology. If you doubt the existence of this dilemma, review at your leisure the applicable titles in the current literature and ask yourself how you would represent them to a foreman, a manager, or a chief executive of your acquaintance in a business enterprise. Imagine how he might react.

WHAT IS THE SUBJECT OF BUSINESS RESEARCH ?

I suspect that one problem is simply that as researchers we do not often enough ask ourselves the questions : What is a business ?

We are more apt to be content with asking where might this technique be applicable, or what problem is there that I can solve with methods I know, but the businessman does not.

We already know that a business is an institution that should make a profit. We also know it should supply customers with values in excess of purchase price or, perhaps better, actual costs, to mention two out of a long list of things a business should do and be according to the experts. There is, thus, a tendency to pursue the obvious.

Now a curious and seldom mentioned fact is that the typical industrial business usually consists of more than one person. In fact if such a business is of any size at all, it usually consists of many people. Although large numbers of these may be direct workers, there are also managers, general, functional and subfunctional, and also professional specialists. These people must find personal satisfaction in their work through voluntarily and knowledgeably contributing of their skills, otherwise there is no profit, no added value, in fact, no business.

In short, a very basic problem of a business is that of somehow obtaining integrated joint performance among people on a voluntary basis and in such a manner that they also get personal satisfaction.

If this is not achieved, the maximum of efficiency in controlling inventory, scheduling production, allocating facilities or deploying salesmen or any other single activity is without particular meaning.

It seems safe to say that until such integrated joint performance reaches an appropriate level of effectiveness, the need for applying more systematic method to other situations in a business may well be obscured to most of its people.

This is not to say that an inventory situation, for example, is not apparent. The physical existence of an inventory may be an observable fact. Its effects on profit may also be apparent in the historical records prepared by the accountants. The solution will, however, likely appear to lie, for example, in the belief that :

1. The sales force is selling the wrong things.
2. Customers' needs are not well enough known.
3. The engineers always try to design something a little different.
4. The stockroom supervisor is doing a sloppy job.
5. Volume of sales is off.

It would seem likely that if such beliefs are held by people, they will not be receptive to a solution involving the theory of Queues or the theory of Markov processes. And it is entirely possible that under the circumstances, they may be right. It is entirely possible that there may be more basic problems — those of acquiring new beliefs and understanding of the business by its people.

Now I must admit that it is always possible to put in place a particular solution by command. This approach usually results in a return to the old ways as soon as the pressure is off. This is not a result of maliciousness, but a result of a lack of understanding of need, purpose and process.

In the year 1960, I find it rather difficult to visualize a stockroom foreman seeing the need for or having either much faith in or understanding of Markov processes, for example. Yet, somehow it is that foreman in concert with others in the business who will make that situation workable.

What has just been noted about the inventory problem would not be true if the stockroom involved could be completely mechanized. The mode of solution could then be designed directly into the machines.

However in 1960 there is little evidence of such a mechanized inventory being feasible, of construction on the one hand. On the other hand, there would still remain the problem of getting people in the business to accept such a solution as :

1. Fitting into the rest of the process of integrated joint performance.

2. Of being applicable over a sufficient time span to pay off in a situation which is not static but always "becoming".

So one comes right back to people and their understanding again.

ONE APPROACH TO MAKING BUSINESS RESEARCH EFFECTIVE

A few minutes back, the curious fact was noted that the typical industrial business usually consists of more than one person.

It was also noted at that time that a very fundamental problem is to achieve integrated joint performance voluntarily through these people with their differing aptitudes, skills and personal objectives.

The existence of this problem is always obscured by the fact that any business, if it becomes successful, achieves to some threshold level of effective joint performance among its people. It is tacitly assumed that this will occur in time or that good managers will get it in time. Businesses that do not get it in some measure just fade away, and there is no evidence remaining.

Yet when one is a customer, it could be apparent, if one were watching for it, how low the level of this threshold really can be.

The emphasis, since world war II, on management by objectives with the related concept of individual responsibility for work and teamwork, as so well exemplified in *New frontiers for professional managers* (1) by Ralph J. Cordiner, is another indication that this problem, though obscured, is nevertheless recognized and real.

In an article entitled "Business objectives and survival needs" (2), Peter Drucker sums up this situation from another viewing point, as follows:

"The management literature is full of discussions of the "problem of the specialist" who sees only his own functional area or of the "problem of the scientist in business" who resents the demand that he subordinates his knowledge to business ends. Yet we will be getting ever more specialized; we will, of necessity, employ more and more highly trained

(1) CORDINER, Ralph J., *New frontiers for professional managers*. McGraw-Hill Book Company, Inc., 1956.

(2) DRUCKER, Peter F., *Business objectives and survival needs: Notes on a discipline of business enterprise*. The Journal of business of the University of Chicago, April, 1958, pp. 82 and 83.

“professionals”. Each of those must be dedicated to his speciality ; yet each must share a common vision and common goals and must voluntarily engage in a common effort. To bring this about is already the most time—and energy—consuming job of management, certainly in our big businesses, and no one I know claims to be able to do it successfully”.

This problem of achieving effective joint performance is probably the most subtle and baffling managerial problem today, in 1960. It will most likely become more pressing and difficult as the complexity of markets, economic relations, social expectations and of physical technology increases over the years.

It also seems likely the natural route for introducing business research as a living activity within an industrial business — something managers and key functional individual contributors will use — *is* contributing to the achievement of effective joint performance among people. It can via this route ultimately lead in time to a natural, participative use of research findings in respect to subfunctional areas of a business as, for example, the aforementioned Markov processes with respect to inventory.

The proposal is not that business research alone will solve this problem of effective joint performance. The proposal is that business research can make a valuable, natural and obvious contribution to its solution — a contribution that can become a part of the working lives of those people in a business who sense, decide, regulate, adjust and innovate.

AN EXAMPLE OF BASIC AREA OF OPPORTUNITY

It seems reasonable to ask the question why this problem of effective joint performance is so difficult and how systematic study might be brought to bear upon it as a living thing.

This question can probably best be answered by an illustration. It should be obvious any single illustrations or small number of illustrations will be specific and, hence, narrow. However, I do hope that the illustration will convey some idea of the general opportunity, as well as the general nature of this fundamental problem which must be attacked if the subject matter of research is the industrial business.

In certain types of equipment businesses, for example, it is not unusual to have listed for sale as many as 20,000 models of product. These models may differ in size, in rating, in particular

features, in detailed performance specification. Although these models of product may all be motors or pumps, for example, and in a sense related the difference between the largest and the smallest model in the product line might be quite marked.

It is also typical of such a business that it has many customers, but that its market nevertheless is not a mass market. The number of customers might be of the order of 5,000 — just enough to be hard to keep track of way down inside the business operations themselves. Some will be large, some small, some will order regularly, others only once in a while.

Of the models listed, some few will likely be in large and regular demand by many customers. The balance will be ordered more or less sporadically, sometimes in large quantities by a particularly good customer, at other times by dribbles. In short, the product mix will be continually shifting. Yet it can be regulated to some degree with appropriate sales effort or with adequate warehousing of finished goods. There is also likely to be a certain percentage of special designs to meet peculiar customer needs at all times. As if this were not enough, improvements are continually being made, most of them of a minor character either by the engineers in this business or by competitors.

Out of this kind of situation, some order has to be made. That order has to be made by and understood operationally in 1960 by people if effective joint performance is to be obtained. It is difficult to visualize it being made in some ivory tower and then imposed effectively, except in some Orwellian *1984*.

It should be apparent that there are at least three opportunities and related risks at this point :

1. The type of order that is achieved will have an important bearing on performance and profits.
2. Once that order is established, the effectiveness with which it is carried through on a day-by-day basis will also have a bearing on performance and profits.
3. The character and timing of purposeful change in the type of order as circumstances change will also be important.

The fundamental difficulty here, as you no doubt surmise, is complexity — truly recognizing it, knowing it can be treated with and, hence, understood, reasoned about and communicated about. Getting this idea across and doing something about it, is probably the first step that has to be taken.

Basic opportunities for business research in 1960 might lie in tackling fundamental questions like the following, in this business of 20,000 product models :

1. How organize information about the 20,000 models — all of which are not necessarily designed or modified at the same time, so that engineers can see in total the possible interrelations among models that do or could exist ?

2. How convey information about mix of orders for product received so that the patterns in this flux of events can be discerned and each order does not need to be considered by itself ?

3. How convey engineering and sales information about such a product line to manufacturing people so that they can devise an effective plan for manufacture ?

4. How convey sufficient detail about both engineering and sales, yet in appropriate form so that not an optimal but at least an appropriate shop can be designed ?

5. How organize information so that difficulties encountered in shop design can be conveyed back to engineers with reasonable expectation that through understanding appropriate product line design, modifications can be made to care for such difficulties where possible ?

6. How organize cost information about such a line so that the effect of model mix on cost may be seen and anticipated, particularly by marketing people ?

7. How organize incoming order information for such a business so that it can be quickly interpreted en masse into designs, shop operations, materials requirement, labor load, rather than by individual order ?

8. How organize information so that a sales plan expressed in units of product may be interpreted and evaluated in similar terms ?

These might be considered as illustrations of the kinds of information needed to establish order.

Lest this illustration be viewed too simply, it should be noted that there are basically two problems in each illustration :

1. Making clear the organization of the information and its relevance to the situation so that people may use the information for reasoning and deciding.

2. The mechanics of processing data so that the information in appropriate and understood form becomes available.

It is not just the latter processing data. In fact, the more important task is the former as this permits people to understand more completely and, hence, to reason about situations with the information made available. It is also only by this route that the need for more research will be perceived by them.

In addition to questions like those just cited that are concerned with establishing some kind of orderly approach, there are those that are concerned with actually operating in terms of that approach day-by-day. These might include the following, as examples :

1. How provide the individual engineer in such a business with information that will make it possible for him to design an individual model of product at some time when an order is received so that it will be consistent with an over-all operational plan for the business ?

2. How provide a sales manager or a salesman with information appropriately organized so that he can take an order on the spur of the moment and make commitments to his customer with respect to it, but with assurance that these are consistent with over-all operating plans at a particular time ?

3. How provide a sales manager with information so that he can determine that the mix of models of orders being received is consistent with profitability goals or the mix that then needs to be obtained in the future if such goals are to be met ?

4. What kind of measurements of demand for products should a warehouse supervisor have, to determine which models of product to stock and when and how much to change stock levels ?

5. What criteria should foremen of component manufacturing areas have available for judging and regulating performance in terms of final assembly operations ?

These two sets of illustrative questions arising out of needs for information in a business that has 5,000 customers and lists 20,000 models of product for sale were cited, if you will recall, to in part answer the question : *Why is the achievement of effective joint performance so difficult ?* The illustrations were intended to make apparent that there is immediately involved matter of complexity and in a very personal, individual way for those people involved in a business.

It is believed that at this time these kinds of questions go largely unanswered but that systematic study could bring appropriately organized information.

Since people within a business are regularly wrestling with ques-

tions of this type, research work directed toward providing information about them, so that people may use it in sensing, reasoning and deciding could make business research a vital and living activity in a business.

CONCLUSION

In summing up, I would like to go back to my initial assumption. This assumption is that for the foreseeable future, most industrial businesses will be pretty much constituted as they are now. They will be loosely coupled systems in which many people will continue to play a major role in sensing, deciding, regulating, adjusting and innovating. They will, moreover, have legitimate authority and responsibility for so doing.

My point is that if business research is to have real and growing impact in such businesses, it must be by and through these people. It must be shown to aid them in *their* reasoning processes.

Developing the use of applied-business research is *not merely doing research*, but it involves a peculiar form of teaching since people are involved. This teaching takes the form of doing research on the gross or fundamental situations first so that people can reason more effectively about them. More detailed research work may come later.

One example of such a situation is that of getting effective joint performance, of depicting the patterns needed to answer questions like the baker's dozen cited earlier in the description of a business selling 20,000 models.

When it becomes apparent that these kinds of situations can be treated with, the opportunities for more tightly coupled systems and the need for refinements will become apparent but not likely before.

I have alluded earlier to the theory of Markov processes applied to the inventory problem. It is not my intent to "pick on" the theory of Markov processes, but I would like to allude to this case again.

I am of the opinion that this means for understanding the ordering-inventory problem is applicable in many instances. However, if a business is at the point where it has not yet measured demands, and only measures and thinks of orders received, and then only as an arithmetic sum, and probably only in aggregated dollars over many items — not only will it not be known to people in the business if this approach is applicable — the approach

itself will likely be meaningless to them. Something else must come first — namely, thinking about and measuring demands as well as order inputs — *both* in some orderly and appropriate manner that displays more about the phenomena than a dollar sum or a sum of units of product for some period.

It will also have been necessary to have recognized that different ordering policies applied on the factory by a number of warehouses do have different effects upon factory operations — some for better, some for worse.

An industrial business is still a system of people and ideas, where “need to”, “want to” and “ought to” as applied to those people and by them is of great significance.

Business research, then, should most likely be viewed as a step-by-step process of mutual learning and mutual reasoning, and to be most effective, it should probably begin with basic problems like those cited earlier — working out from these to the more refined and detailed.

